Construction of Multiple Access Channel Codes Based on Hash Property

Jun Muramatsu and Shigeki Miyake

Abstract

The aim of this paper is to introduce the construction of codes for a general discrete stationary memoryless multiple access channel based on the the notion of the hash property. Since an ensemble of sparse matrices has a hash property, we can use sparse matrices for code construction. Our approach has a potential advantage compared to the conventional random coding because it is expected that we can use some approximation algorithms by using the sparse structure of codes.

Index Terms

Shannon theory, hash property, linear codes, LDPC codes, sparse matrix, minimum-divergence encoding/decoding, multiple access channel.

I. Introduction

This paper describes the construction of multiple access channel codes. In a multiple access channel, two or more senders send messages to a common receiver. The capacity region has been derived in [1][15] for a scenario where two senders have different private messages but no common message to be sent. This work has been extended in [26] to a scenario where two senders have different private messages and a common message to be sent. The capacity region for two or more senders has been described in [6, Section 15.3.5][10, Chapter 4] in which there is no common message. In [12], the capacity region has been derived for a general multiple access channel in which two or more senders have messages common to some users. Applications of Low Density Parity Check (LDPC) codes to a multiple access channel have been introduced in [3][16][17]. Furthermore, there are many theoretical/experimental studies regarding the construction of multiple access channel codes by using LDPC codes, e.g. [2][25]. It should be noted that they assumed channel noises to be additive.

A contribution of this paper is to construct codes based on the notion of the hash property [22][21], which is a stronger version of that introduced in [19][20]. Another contribution is to construct codes by using sparse matrices for a general discrete memoryless multiple access channel including asymmetric one. We construct codes for the following scenarios:

- Two or more senders have different private messages (Fig.1) [10, Theorem 5 in Chapter 4][6, Section 15.3.5],

J. Muramatsu is with NTT Communication Science Laboratories, NTT Corporation, 2-4, Hikaridai, Seika-cho, Soraku-gun, Kyoto 619-0237, Japan (E-mail: muramatsu.jun@lab.ntt.co.jp). S. Miyake is with NTT Network Innovation Laboratories, NTT Corporation, Hikarinooka 1-1, Yokosuka-shi, Kanagawa 239-0847, Japan (E-mail: miyake.shigeki@lab.ntt.co.jp). This paper has been presented in part at [24] and submitted to IEEE Transactions on Information Theory.

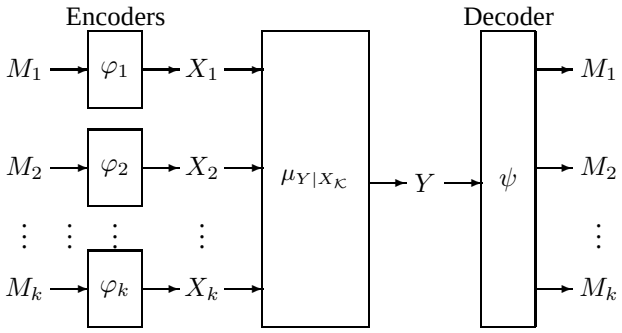

Fig. 1. Multiple Access Channel Coding: Private Messages

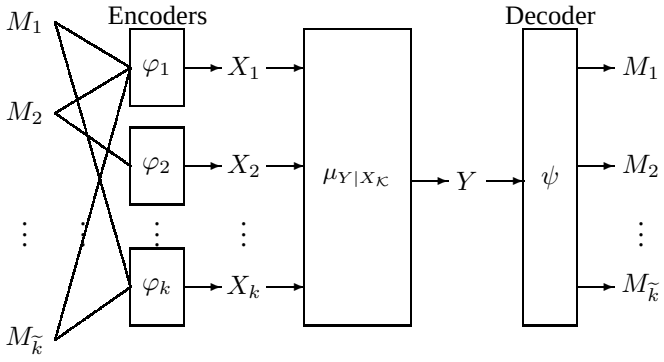

Fig. 2. Multiple Access Channel Coding: Multiple Common Messages

- Two or more senders have messages common to senders (Fig.2) [12], and
- Two senders have different private messages and a common message (Fig.3) [26],

where additive channel noises are not assumed. It should be noted that the first scenario includes two-sender scenario studied in [1][15]. The last scenario is included in the second scenario but we will discuss it separately because their code constructions are different. The proof of all the theorems is based on the notion of the hash property. It is an extension of the ensemble of the random bin coding [5], the ensembles of linear matrices [7], the universal class of hash functions [9], and the ensemble of sparse matrices [18]. We use two lemmas, one related to 'saturation property[1]' (if the number of items is greater than the number of bins then there is an assignment such that every bin contains at least one item) and the other related to 'collision-resistance property[1]' (if the number of bins is greater than the number of items then there is an assignment such that every bin contains at most one item) proved in [19][21], where the lemma related to the 'collision-resistance property' is extended from a single domain to multiple domains. They are reviewed in Section IV. The saturation property is used to analyze the average encoding error and the extended collision-resistance property is used to analyze the average decoding error. It should be noted that the functions need not be linear for the hash property but it is expected that the space and time complexity of codes can be reduced compared with conventional constructions by

[1]In [19], they were called 'saturating property' and 'collision-resistant property,' respectively. We changed these terms following the suggestion of Prof. T.S. Han.

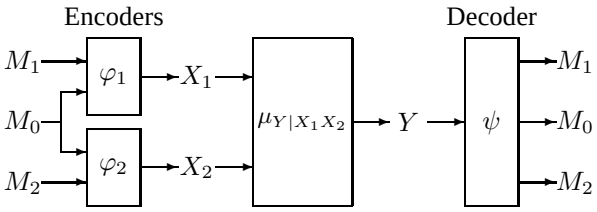

Fig. 3. Two-user Multiple Access Channel Coding: Private and Common Messages

assuming the linearity of functions. In fact, we can construct codes by using sparse matrices (with logarithmic column degree) because an ensemble of sparse matrices has a hash property [21]. Although the implementation of encoding and decoding operations of our approach is still intractable, our approach has a potential advantage compared to the conventional random coding presented in [6][7][10][12][26] because it is expected[2] that we can use some approximation algorithms such as the sum-product algorithm [14] and the linear programing algorithm [11] for encoding and decoding operations in the proposed code with sparse matrix.

II. Definitions and Notations

Throughout this paper, we use the following definitions and notations. The cardinality of a set $\mathcal{U}$ is denoted by $|\mathcal{U}|$, $\mathcal{U}^c$ denotes the complement of $\mathcal{U}$, and $\mathcal{U} \setminus \mathcal{V} \equiv \mathcal{U} \cap \mathcal{V}^c$ denotes the set difference.

Column vectors and sequences are denoted in boldface. Let $A\boldsymbol{u}$ denote a value taken by a function $A : \mathcal{U}^n \to \overline{\mathcal{U}}$ at $\boldsymbol{u} \in \mathcal{U}^n$, where $\mathcal{U}^n$ is the domain of the function and $\overline{\mathcal{U}}$ is the range of function. It should be noted that A may be nonlinear. When A is a linear function expressed by an $l \times n$ matrix, we assume that $\mathcal{U} \equiv \mathrm{GF}(q)$ is a finite field and the range of functions is $\mathcal{U}^l$. For a set $\mathcal{A}$ of functions, let $\mathrm{Im}\mathcal{A}$ be defined as

$$\mathrm{Im}\mathcal{A} \equiv \bigcup_{A \in \mathcal{A}} \{A\boldsymbol{u} : \boldsymbol{u} \in \mathcal{U}^n\}.$$

We define a set $\mathcal{C}_A(\boldsymbol{a})$ as

$$\mathcal{C}_A(\boldsymbol{a}) \equiv \{\boldsymbol{u} : A\boldsymbol{u} = \boldsymbol{a}\}$$

In the context of linear codes, $\mathcal{C}_A(\boldsymbol{a})$ is called a coset determined by $\boldsymbol{a}$. The random variables of a function A and a vector $\boldsymbol{a} \in \mathrm{Im}\mathcal{A}$ are denoted by the sans serif letters A and $\mathbf{a}$, respectively. It should be noted that some random variables are denoted by the Roman letter (e.g. M, U, V, X, Y) which does not represent a function.

For random variables U and V, let μ_U be the probability distribution and $\mu_{U|V}$ be the conditional probability distribution. Then the entropy $H(U)$, the conditional entropy $H(U|V)$, and the mutual information $I(U;V)$ are defined as

$$H(U) \equiv \sum_u \mu_U(u) \log \frac{1}{\mu_U(u)}$$

$$H(U|V) \equiv \sum_{u,v} \mu_{U|V}(u|v)\mu_V(v) \log \frac{1}{\mu_{U|V}(u|v)}$$

[2]In fact, the direct application of [11][14] does not provide good performance. Implementation of our approach is left for a future challenge.

$$I(U;V) \equiv H(U) - H(U|V)$$

where we assume a logarithm with base 2 when the subscript of $\log$ is omitted. For random variables U, V, and W, let $I(U;V|W)$ be the conditional mutual information defined as

$$I(U;V|W) \equiv H(U|W) - H(U|V,W).$$

For $\boldsymbol{u} \in \mathcal{U}^n$ and $\boldsymbol{v} \in \mathcal{V}^n$, let $\nu_{\boldsymbol{u}}$ and $\nu_{\boldsymbol{u}|\boldsymbol{v}}$ be the empirical distributions defined as

$$\nu_{\boldsymbol{u}}(u) \equiv \frac{|\{i \in \{1,\ldots,n\} : u_i = u\}|}{n} \tag{1}$$

$$\nu_{\boldsymbol{u}|\boldsymbol{v}}(u|v) \equiv \frac{\nu_{\boldsymbol{u}\boldsymbol{v}}(u,v)}{\nu_{\boldsymbol{v}}(v)} \quad \text{for } v \in \mathcal{V} \text{ s.t. } \nu_{\boldsymbol{v}}(v) > 0, \tag{2}$$

where we use the relation $\nu_{\boldsymbol{u}|\boldsymbol{v}}(u|v)\nu_{\boldsymbol{v}}(v) = \nu_{\boldsymbol{u}\boldsymbol{v}}(u,v)$ even when $\nu_{\boldsymbol{v}}(v) = 0$. Let p and p' be probability distributions on the same set $\mathcal{U}$ and let q and q' be conditional probability distributions on the same set $\mathcal{V}$. Then divergence $D(p\|p')$ and conditional divergence $D(q\|q'|p)$ are defined as

$$D(p \parallel p') \equiv \sum_{u\in\mathcal{U}} p(u) \log \frac{p(u)}{p'(u)}$$

$$D(q \parallel q'|p) \equiv \sum_{u\in\mathcal{U}} p(v) \sum_{v\in\mathcal{V}} q(v|u) \log \frac{q(v|u)}{q'(v|u)}.$$

For the proof of theorems, we use the method of type developed in [8], where we use the definition of a typical set introduced in [19][27]. A set of typical sequences $\mathcal{T}_{U,\gamma}$ and a set of conditionally typical sequences $\mathcal{T}_{U|V,\gamma}(\boldsymbol{v})$ are defined as

$$\mathcal{T}_{U,\gamma} \equiv \{\boldsymbol{u} : D(\nu_{\boldsymbol{u}}\|\mu_U) < \gamma\}$$

$$\mathcal{T}_{U|V,\gamma}(\boldsymbol{v}) \equiv \left\{\boldsymbol{u} : D(\nu_{\boldsymbol{u}|\boldsymbol{v}}\|\mu_{U|V}|\nu_{\boldsymbol{v}}) < \gamma\right\},$$

respectively. For $\boldsymbol{u} \in \mathcal{X}^n$, $\boldsymbol{v} \in \mathcal{V}^n$, the entropy $H(\boldsymbol{u})$, and the conditional entropy $H(\boldsymbol{u}|\boldsymbol{v})$ are defined as

$$H(\boldsymbol{u}) \equiv \sum_{u} \nu_{\boldsymbol{u}}(u) \log \frac{1}{\nu_{\boldsymbol{u}}(u)}$$

$$H(\boldsymbol{u}|\boldsymbol{v}) \equiv \sum_{u,v} \nu_{\boldsymbol{u}|\boldsymbol{v}}(u|v)\nu_{\boldsymbol{v}}(v) \log \frac{1}{\nu_{\boldsymbol{u}|\boldsymbol{v}}(u|v)},$$

where $\nu_{\boldsymbol{u}}$ and $\nu_{\boldsymbol{u}|\boldsymbol{v}}$ are defined as (1) and (2), respectively. For $\gamma, \gamma' > 0$, we define

$$\lambda_{\mathcal{U}} \equiv \frac{|\mathcal{U}| \log(n+1)}{n} \tag{3}$$

$$\iota_{\mathcal{U}}(\gamma) \equiv -\sqrt{2\gamma} \log \frac{\sqrt{2\gamma}}{|\mathcal{U}|} \tag{4}$$

$$\iota_{\mathcal{U}|\mathcal{V}}(\gamma'|\gamma) \equiv -\sqrt{2\gamma'} \log \frac{\sqrt{2\gamma'}}{|\mathcal{U}||\mathcal{V}|} + \sqrt{2\gamma} \log |\mathcal{U}| \tag{5}$$

$$\eta_{\mathcal{U}}(\gamma) \equiv -\sqrt{2\gamma} \log \frac{\sqrt{2\gamma}}{|\mathcal{U}|} + \frac{|\mathcal{U}| \log(n+1)}{n} \tag{6}$$

$$\eta_{\mathcal{U}|\mathcal{V}}(\gamma'|\gamma) \equiv -\sqrt{2\gamma'} \log \frac{\sqrt{2\gamma'}}{|\mathcal{U}||\mathcal{V}|} + \sqrt{2\gamma} \log |\mathcal{U}| + \frac{|\mathcal{U}||\mathcal{V}| \log(n+1)}{n}, \tag{7}$$

which comes from lemmas in Appendix B regarding the method of types. It should be noted here that the product set $\mathcal{U} \times \mathcal{V}$ is denoted by $\mathcal{U}\mathcal{V}$ when it appears in the subscript of these functions.

For a mathematical statement S, we define $\chi(S)$ as

$$\chi(S) \equiv \begin{cases} 1, & \text{if } S \text{ is true} \\ 0, & \text{if } S \text{ is false.} \end{cases}$$

III. Formal Description of Problems and Known Results

In this section, we review the problems of multiple access channel coding and results regarding achievable regions.

A multiple access channel has k inputs and 1 output. Let $\mathcal{K}$ be an index set of the channel inputs, where $k \equiv |\mathcal{K}|$. Then the channel is characterized by the conditional probability distribution $\mu_{Y|X_\mathcal{K}}$, where $X_\mathcal{K} \equiv \{X_j\}_{j\in\mathcal{K}}$ is a k-tuple of random variables corresponding to the inputs and Y is a random variable corresponding to the output. Let $\mathcal{X}_j$ be the alphabet of the j-th channel input and $\mathcal{Y}$ be the alphabet of the channel output.

In the following, we review some coding scenarios that will be discussed in subsequent sections. Let $\widetilde{\mathcal{K}}$ be an index set of messages and $\widetilde{k} \equiv |\widetilde{\mathcal{K}}|$. For each $i \in \widetilde{\mathcal{K}}$, let $\mathcal{M}_i$ be the alphabet of the i-th message and M_i be the random variable corresponding to the i-th message, where we assume that the probability distribution of M_i is uniform on $\mathcal{M}_i$ for all $i \in \widetilde{\mathcal{K}}$. We also assume that random variables $\{M_i\}_{i\in\widetilde{\mathcal{K}}}$ are mutually independent. Let $p_{M_{\widetilde{\mathcal{K}}}}$ be the uniform distribution on $\mathcal{M}_{\widetilde{\mathcal{K}}}$. We use the following notations:

$$\begin{aligned} \varphi_\mathcal{K} &\equiv \{\varphi_j\}_{j\in\mathcal{K}} \\ \mathcal{M}_{\widetilde{\mathcal{K}}} &\equiv \bigtimes_{i\in\widetilde{\mathcal{K}}} \mathcal{M}_i \\ \boldsymbol{m}_{\widetilde{\mathcal{K}}} &\equiv \{\boldsymbol{m}_i\}_{i\in\widetilde{\mathcal{K}}}, \quad \text{for given } \boldsymbol{m}_i \in \mathcal{M}_i,\ i \in \widetilde{\mathcal{K}} \\ R_{\widetilde{\mathcal{K}}} &\equiv \{R_i\}_{i\in\widetilde{\mathcal{K}}}. \end{aligned}$$

Let $\mathrm{cl}(\cdot)$ denote the closure of a region and $\mathrm{co}(\cdot)$ denote the closure of the convex hull of a region.

A. Private Messages

In this scenario, we assume that $\widetilde{\mathcal{K}} = \mathcal{K}$ and there are k senders and k independent messages $M_\mathcal{K}$, where the j-th sender has access to the j-th message M_j and there is no common message.

For a given block length n, a multiple access channel code $(\varphi_\mathcal{K}, \psi)$ (Fig.1) is defined by k encoders $\varphi_\mathcal{K}$ and one decoder ψ, where

$$\begin{aligned} \varphi_j &: \mathcal{M}_j \to \mathcal{X}_j^n \quad \text{for each } j \in \mathcal{K} \\ \psi &: \mathcal{Y}^n \to \mathcal{M}_\mathcal{K}. \end{aligned}$$

Then the error probability of the code is defined as

$$\mathrm{Error}(\varphi_\mathcal{K}, \psi) \equiv \sum_{\substack{\boldsymbol{m}_\mathcal{K}\in\mathcal{M}_\mathcal{K} \\ \boldsymbol{y}\in\mathcal{Y}^n}} \mu_{Y|X_\mathcal{K}}(\boldsymbol{y}|\varphi_\mathcal{K}(\boldsymbol{m}_\mathcal{K}))p_{M_\mathcal{K}}(\boldsymbol{m}_\mathcal{K})\chi(\psi(\boldsymbol{y}) \neq \boldsymbol{m}_\mathcal{K}).$$

The rate R_j of the j-th message is defined as

$$R_j \equiv \frac{\log|\mathcal{M}_j|}{n} \quad \text{for each } j \in \mathcal{K}.$$

We call the rate vector $R_{\mathcal{K}}$ *achievable* if for all $\delta > 0$ and all sufficiently large n, there is a code $(\varphi_{\mathcal{K}}, \psi)$ with a rate vector $R_{\mathcal{K}}$ such that

$$\mathrm{Error}(\varphi_{\mathcal{K}}, \psi) < \delta.$$

For a given $\{\mu_{X_j}\}_{j\in\mathcal{K}}$, let $\mathcal{R}(\{\mu_{X_j}\}_{j\in\mathcal{K}})$ be the set of all k-dimensional vectors $R_{\mathcal{K}}$ satisfying

$$0 \le \sum_{j\in\mathcal{J}} R_j < I(X_{\mathcal{J}}; Y|X_{\mathcal{J}^c}) \quad \text{for all } \mathcal{J} \subset \mathcal{K}, \tag{8}$$

where the joint distribution $\mu_{X_{\mathcal{K}}Y}$ of random variable $(X_{\mathcal{K}}, Y)$ is given by

$$\mu_{X_{\mathcal{K}}Y}(x_{\mathcal{K}}, y) \equiv \mu_{Y|X_{\mathcal{K}}}(y|x_{\mathcal{K}}) \left[\prod_{j\in\mathcal{K}} \mu_{X_j}(x_j)\right]. \tag{9}$$

For given μ_U and $\{\mu_{X_j|U}\}_{j\in\mathcal{K}}$, let $\mathcal{R}(\mu_U, \{\mu_{X_j|U}\}_{j\in\mathcal{K}})$ be the set of all k-dimensional vectors $R_{\mathcal{K}}$ satisfying

$$0 \le \sum_{j\in\mathcal{J}} R_j < I(X_{\mathcal{J}}; Y|U, X_{\mathcal{J}^c}) \quad \text{for all } \mathcal{J} \subset \mathcal{K}, \tag{10}$$

where the joint distribution $\mu_{UX_{\mathcal{K}}Y}$ of random variable $(U, X_{\mathcal{K}}, Y)$ is given by

$$\mu_{UX_{\mathcal{K}}Y}(u, x_{\mathcal{K}}, y) \equiv \mu_{Y|X_{\mathcal{K}}}(y|x_{\mathcal{K}}) \left[\prod_{j\in\mathcal{K}} \mu_{X_j|U}(x_j|u)\right] \mu_U(u). \tag{11}$$

Then the achievable region for this scenario is given as described below.

Proposition 1 ([6, Theorem 15.3.6][10, Theorem 4.5]): The achievable region for this scenario is given as

$$\mathrm{co}\left(\bigcup_{\{\mu_{X_j}\}_{j\in\mathcal{K}}} \mathcal{R}(\{\mu_{X_j}\}_{j\in\mathcal{K}})\right), \tag{12}$$

which is equivalent to

$$\bigcup_{\mu_U, \{\mu_{X_j}\}_{j\in\mathcal{K}}} \mathrm{cl}\left(\mathcal{R}(\mu_U, \{\mu_{X_j|U}\}_{j\in\mathcal{K}})\right), \tag{13}$$

where $|\mathcal{U}| \le k$.

Remark 1: It should be noted that this proposition includes the result of [1][15] corresponding to the case of two encoders. The equivalence of the two regions (12) and (13) can be shown from [6, Theorem 15.3.6] and [10, Theorem 4.5] by considering the operational definition of capacity region.

In Section V-A, for a given $R_{\mathcal{K}} \in \mathcal{R}(\mu_U, \{\mu_{X_j|U}\}_{j\in\mathcal{K}})$, we construct a code with the rate vector $R_{\mathcal{K}}$ based on the coded time sharing technique. It should be noted that we can construct a code with a rate vector $R_{\mathcal{K}} \in \mathcal{R}(\{\mu_{X_j}\}_{j\in\mathcal{K}})$ by letting U be a constant, that is, $|\mathcal{U}| = 1$. In fact, $\mathcal{R}(\{\mu_{X_j}\}_{j\in\mathcal{K}}) = \mathcal{R}(\mu_U, \{\mu_{X_j|U}\}_{j\in\mathcal{K}})$ when U is a constant. The achievability of the region (12) with a proposed code can be proved by using the time-sharing argument.

B. Multiple Common Messages

In this scenario, we assume that there are $\widetilde{k}$ independent messages $M_{\widetilde{\mathcal{K}}}$ and k encoders, where the j-th encoder has access to the messages $M_{\widetilde{\mathcal{K}}_j} \equiv \{M_i\}_{i\in\widetilde{\mathcal{K}}_j}$ specified by $\widetilde{\mathcal{K}}_j \subset \widetilde{\mathcal{K}}$ for each $j \in \mathcal{K}$.

For a given block length n, a multiple access channel code $(\varphi_{\mathcal{K}}, \psi)$ (Fig.2) is defined by k encoders $\varphi_{\mathcal{K}} \equiv \{\varphi_j\}_{j\in\mathcal{K}}$ and one decoder ψ, where

$$\varphi_j : \mathcal{M}_{\widetilde{\mathcal{K}}_j} \to \mathcal{X}_j^n \quad \text{for each } j \in \mathcal{K}$$

$$\psi : \mathcal{Y}^n \to \mathcal{M}_{\mathcal{K}}.$$

Then the error probability of the code is defined by

$$\mathrm{Error}(\varphi_{\mathcal{K}}, \psi) \equiv \sum_{\substack{\boldsymbol{m}_{\widetilde{\mathcal{K}}} \in \mathcal{M}_{\widetilde{\mathcal{K}}} \\ \boldsymbol{y} \in \mathcal{Y}^n}} \mu_{Y|X_{\mathcal{K}}}(\boldsymbol{y}|\varphi_{\mathcal{K}}(\boldsymbol{m}_{\widetilde{\mathcal{K}}}))p_{M_{\widetilde{k}}}(\boldsymbol{m}_{\widetilde{\mathcal{K}}})\chi(\psi(\boldsymbol{y}) \neq \boldsymbol{m}_{\widetilde{\mathcal{K}}}).$$

For each $i \in \widetilde{\mathcal{K}}$, the rate R_i of the i-th message is defined by

$$R_i \equiv \frac{\log |\mathcal{M}_i|}{n}.$$

We call the rate vector $R_{\widetilde{\mathcal{K}}}$ *achievable* if for all $\delta > 0$ and all sufficiently large n, there is a code $(\varphi_{\mathcal{K}}, \psi)$ with a rate vector $R_{\widetilde{\mathcal{K}}}$ such that

$$\mathrm{Error}(\varphi_{\mathcal{K}}, \psi) < \delta.$$

For each $i \in \widetilde{\mathcal{K}}$, let $\widetilde{X}_i$ be an auxiliary random variable and $\widetilde{\mathcal{X}}_i$ is the alphabet of $\widetilde{X}_i$. For a given $\{\mu_{\widetilde{X}_i}\}_{i\in\widetilde{\mathcal{K}}}$ and a set $\{f_j\}_{j\in\mathcal{K}}$ of functions

$$f_j : \widetilde{\mathcal{X}}_{\widetilde{\mathcal{K}}_j} \to \mathcal{X}_j \quad \text{for each } j \in \mathcal{K},$$

let[3] $\mathcal{R}_{\mathrm{H}}(\{\mu_{\widetilde{X}_i}\}_{i\in\widetilde{\mathcal{K}}}, \{f_j\}_{j\in\mathcal{K}})$ be the set of all s-dimensional vectors $R_{\widetilde{\mathcal{K}}}$ satisfying

$$0 \leq \sum_{i\in\mathcal{I}} R_i < I(\widetilde{X}_{\mathcal{I}}; Y|\widetilde{X}_{\mathcal{I}^c}) \quad \text{for all } \mathcal{I} \subset \widetilde{\mathcal{K}}, \tag{14}$$

where the joint distribution $\mu_{\widetilde{X}_{\widetilde{\mathcal{K}}}X_{\mathcal{K}}Y}$ of random variable $(\widetilde{X}_{\widetilde{\mathcal{K}}}, X_{\mathcal{K}}, Y)$ is given by

$$\mu_{\widetilde{X}_{\widetilde{\mathcal{K}}}X_{\mathcal{K}}Y}(\widetilde{x}_{\widetilde{\mathcal{K}}}, x_{\mathcal{K}}, y) \equiv \mu_{Y|X_{\mathcal{K}}}(y|x_{\mathcal{K}}) \left[\prod_{j\in\mathcal{K}} \chi(f_j(\widetilde{x}_{\widetilde{\mathcal{K}}_j}) = x_j)\right] \left[\prod_{i\in\widetilde{\mathcal{K}}} \mu_{\widetilde{X}_i}(\widetilde{x}_i)\right].$$

Then the achievable region for this scenario is given as described below.

Proposition 2 ([12, Theorem 4.1]): The achievable region for this scenario is given as

$$\mathrm{co}\left(\bigcup_{\{\mu_{\widetilde{X}_i}\}_{i\in\widetilde{\mathcal{K}}}, \{f_j\}_{j\in\mathcal{K}}} \mathcal{R}_{\mathrm{H}}(\{\mu_{\widetilde{X}_i}\}_{i\in\widetilde{\mathcal{K}}}, \{f_j\}_{j\in\mathcal{K}})\right), \tag{15}$$

where

$$|\widetilde{\mathcal{X}}_i| \leq |\widetilde{\mathcal{K}}| + \prod_{\substack{j\in\mathcal{K}: \\ i\in\widetilde{\mathcal{K}}_j}} |\mathcal{X}_j| \quad \text{for all } i \in \widetilde{\mathcal{K}}.$$

In Section V-B, for a given $\mathcal{R}_{\mathrm{H}}(\{\mu_{\widetilde{X}_i}\}_{i\in\widetilde{\mathcal{K}}}, \{f_j\}_{j\in\mathcal{K}})$, we construct a code with the rate vector $R_{\widetilde{\mathcal{K}}}$. The achievability of region (15) with a proposed code can be proved by using the time-sharing argument.

In the following, let us consider a scenario (Fig.3) in which one of two senders has access to messages M_0 and M_1 and another sender has access to messages M_0 and M_2, where M_0 denotes a common message. It is a special case of the above scenario, where $\mathcal{K} \equiv \{1, 2\}$, $\widetilde{\mathcal{K}} \equiv \{0, 1, 2\}$, $\widetilde{\mathcal{K}}_1 \equiv \{0, 1\}$, and $\widetilde{\mathcal{K}}_2 \equiv \{0, 2\}$.

Let R_0 be the encoding rate of the common message and R_1 and R_2 be the encoding rate of the private message of the respective encoders. Let[4] $\mathcal{R}_{\mathrm{SW}}(\mu_{X_0}, \mu_{X_1|X_0}, \mu_{X_2|X_0})$ be the set of all (R_0, R_1, R_2) satisfying

$$R_0 \geq 0 \tag{16}$$

[3]The subscript 'H' comes from the author Han of [12].

[4]The subscript 'SW' comes from the authors Slepian and Wolf of [26].

$$0 \leq R_1 < I(X_1; Y|X_0, X_2) \tag{17}$$

$$0 \leq R_2 < I(X_2; Y|X_0, X_1) \tag{18}$$

$$R_1 + R_2 < I(X_1, X_2; Y|X_0) \tag{19}$$

$$R_0 + R_1 + R_2 < I(X_1, X_2; Y) \tag{20}$$

where the joint distribution $\mu_{X_0X_1X_2Y}$ of random variables (X_0, X_1, X_2, Y) is given by

$$\mu_{X_0X_1X_2Y}(x_0, x_1, x_2, y) \equiv \mu_{Y|X_1X_2}(y|x_1, x_2)\mu_{X_1|X_0}(x_1|x_0)\mu_{X_2|X_0}(x_2|x_0)\mu_{X_0}(x_0). \tag{21}$$

It should be noted that (21) implies the fact that the right hand side of (20) is equal to $I(X_1X_2X_0; Y)$. Then, the rate region is given as described below.

Proposition 3 ([26]): For the scenario in which two receivers have access to their private message and a common message, the achievable region is given as

$$\mathrm{co}\left(\bigcup_{\mu_{X_0}, \mu_{X_1|X_0}, \mu_{X_2|X_0}} \mathcal{R}_{\mathrm{SW}}(\mu_{X_0}, \mu_{X_1|X_0}, \mu_{X_2|X_0})\right), \tag{22}$$

where

$$|\mathcal{X}_0| \leq \min\{|\mathcal{Y}| + 3, |\mathcal{X}_1||\mathcal{X}_2| + 2\}.$$

Remark 2: It should be noted that region (22) is equivalent to the region obtained from (15). This has been proven in [12].

In Section V-C, for given $(R_0, R_1, R_2) \in \mathcal{R}_{\mathrm{SW}}(\mu_{X_0}, \mu_{X_1|X_0}, \mu_{X_2|X_0})$, we construct a code with the rate vector (R_0, R_1, R_2). The construction is a typical example of the superposition coding introduced in [26] based on the hash property, and it is different from the construction presented in Section V-B. The achievability of region (22) with a proposed code can be proved by using the time-sharing argument.

IV. $(\boldsymbol{\alpha}, \boldsymbol{\beta})$-HASH PROPERTY

In this section, we introduce the hash property first introduced in [22][21] and its implications. This notion is used for the proof of theorems.

A. Formal Definition

Here, we introduce the hash property for an ensemble of functions. It has been introduced in [22][21] and requires stronger conditions than those introduced in [19].

Definition 1 ([21][22]): Let $\boldsymbol{\mathcal{A}} \equiv \{\mathcal{A}^{(n)}\}_{n=1}^{\infty}$ be a sequence of sets such that $\mathcal{A}^{(n)}$ is a set of functions $A : \mathcal{U}^n \to \mathrm{Im}\mathcal{A}^{(n)}$. For a probability distribution $p_{\mathsf{A},n}$ on $\mathcal{A}^{(n)}$, we call a sequence $(\boldsymbol{\mathcal{A}}, \boldsymbol{p}_{\mathsf{A}}) \equiv \{(\mathcal{A}^{(n)}, p_{\mathsf{A},n})\}_{n=1}^{\infty}$ an *ensemble*. Then, $(\boldsymbol{\mathcal{A}}, \boldsymbol{p}_{\mathsf{A}})$ has a $(\boldsymbol{\alpha}_{\mathsf{A}}, \boldsymbol{\beta}_{\mathsf{A}})$-*hash property*[5] (or simply *hash property*) if there are two sequences $\boldsymbol{\alpha}_{\mathsf{A}} \equiv \{\alpha_{\mathsf{A}}(n)\}_{n=1}^{\infty}$ and $\boldsymbol{\beta}_{\mathsf{A}} \equiv \{\beta_{\mathsf{A}}(n)\}_{n=1}^{\infty}$, which depend on $\{p_{\mathsf{A},n}\}_{n=1}^{\infty}$, such that

$$\lim_{n\to\infty} \alpha_{\mathsf{A}}(n) = 1 \tag{H1}$$

$$\lim_{n\to\infty} \beta_{\mathsf{A}}(n) = 0 \tag{H2}$$

[5]In [21][22][23][24], it is called the 'strong hash property.' Throughout this paper, we call it simply the 'hash property.'

and

$$\sum_{\substack{\boldsymbol{u}' \in \mathcal{U}^n \setminus \{\boldsymbol{u}\}: \\ p_{\mathsf{A},n}(\{A: A\boldsymbol{u}=A\boldsymbol{u}'\}) > \frac{\alpha_{\mathsf{A}}(n)}{|\mathrm{Im}\mathcal{A}_n|}}} p_{\mathsf{A},n}\left(\{A : A\boldsymbol{u} = A\boldsymbol{u}'\}\right) \le \beta_{\mathsf{A}}(n) \tag{H3}$$

for any n and $\boldsymbol{u} \in \mathcal{U}^n$. Throughout this paper, we omit the dependence of $\mathcal{A}$, p_{A}, α_{A} and β_{A} on n.

Remark 3: In [19][22], an ensemble is required to satisfy the condition

$$\lim_{n\to\infty} \frac{1}{n} \log \frac{|\overline{\mathcal{U}}^{(n)}|}{|\mathrm{Im}\mathcal{A}^{(n)}|} = 0,$$

where $\overline{\mathcal{U}}^{(n)}$ is the range of functions. This condition is omitted because it is unnecessary for the results reported in this paper.

Let us remark on the condition (H3). This condition requires the sum of the collision probabilities $p_{\mathsf{A}}\left(\{A : A\boldsymbol{u} = A\boldsymbol{u}'\}\right)$, which is greater than $\alpha_{\mathsf{A}}/|\mathrm{Im}\mathcal{A}|$, to be bounded by β_{A}, where the sum is taken over all $\boldsymbol{u}'$ except $\boldsymbol{u}$. For an ensemble of sparse matrices, α_{A} represents the difference between $(\boldsymbol{\mathcal{A}}, \boldsymbol{p}_{\mathsf{A}})$ and the ensemble of all linear matrices with uniform distribution, and β_{A} represents the upper bound of the probability that the set $\{\boldsymbol{u} \in \mathcal{U}^n : A\boldsymbol{u} = 0\}$, which is called a code in the context of linear codes, has low weight vectors. It should be noted that this condition implies

$$\sum_{\substack{\boldsymbol{u}\in\mathcal{T} \\ \boldsymbol{u}'\in\mathcal{T}'}} p_{\mathsf{A}}\left(\{A : A\boldsymbol{u} = A\boldsymbol{u}'\}\right) \le |\mathcal{T} \cap \mathcal{T}'| + \frac{|\mathcal{T}||\mathcal{T}'|\alpha_{\mathsf{A}}}{|\mathrm{Im}\mathcal{A}|} + \min\{|\mathcal{T}|, |\mathcal{T}'|\}\beta_{\mathsf{A}} \tag{H3'}$$

for any $\mathcal{T}, \mathcal{T}' \subset \mathcal{U}^n$, which is introduced in [19]. The stronger condition (H3) is required for Lemmas 1 and 4, which will appear later.

It should be noted that when $\mathcal{A}$ is a two-universal class of hash functions [9] and p_{A} is the uniform distribution on $\mathcal{A}$, then $(\boldsymbol{\mathcal{A}}, \boldsymbol{p}_{\mathsf{A}})$ has a $(\mathbf{1}, \mathbf{0})$-hash property, where random bin coding [5] and the set of all linear functions [7] are examples of the two-universal class of hash functions. It is proved in [21, Section III-B] that an ensemble of sparse matrices has a hash property. From this fact, this ensemble of sparse matrices can be applied to all results in this paper. This implies that all proposed codes can be constructed by using sparse matrices.

We have the following lemma, where it is unnecessary to assume the linearity of functions assumed in [19][20]. It is one of the advantages of introducing a stronger version of the hash property.

Lemma 1 ([21, Lemma 4]): Let $(\boldsymbol{\mathcal{A}}, \boldsymbol{p}_{\mathsf{A}})$ and $(\boldsymbol{\mathcal{A}}', \boldsymbol{p}_{\mathsf{A}'})$ be ensembles satisfying a $(\boldsymbol{\alpha}_{\mathsf{A}}, \boldsymbol{\beta}_{\mathsf{A}})$-hash property and a $(\boldsymbol{\alpha}_{\mathsf{A}'}, \boldsymbol{\beta}_{\mathsf{A}'})$-hash property, respectively. Let $\mathcal{A} \in \boldsymbol{\mathcal{A}}$ (resp. $\mathcal{A}' \in \boldsymbol{\mathcal{A}}'$) be a set of functions $A : \mathcal{U}^n \to \mathrm{Im}\mathcal{A}$ (resp. $A' : \mathcal{U}^n \to \mathrm{Im}\mathcal{A}'$). Let $\widehat{\mathcal{A}} \equiv \mathcal{A} \times \mathcal{A}'$ and $\widehat{A} \equiv (A, A') \in \widehat{\mathcal{A}}$ defined as

$$\widehat{A}\boldsymbol{u} \equiv (A\boldsymbol{u}, A'\boldsymbol{u}) \quad \text{for each } \widehat{A} \in \widehat{\mathcal{A}},\ \boldsymbol{u} \in \mathcal{U}^n.$$

Let $p_{\widehat{\mathsf{A}}}$ be a joint distribution on $\widehat{\mathcal{A}}$ defined as

$$p_{\widehat{\mathsf{A}}}(A, A') \equiv p_{\mathsf{A}}(A) p_{\mathsf{A}'}(A').$$

Then the ensemble $(\widehat{\boldsymbol{\mathcal{A}}}, \boldsymbol{p}_{\widehat{\mathsf{A}}})$ has a $(\boldsymbol{\alpha}_{\widehat{\mathsf{A}}}, \boldsymbol{\beta}_{\widehat{\mathsf{A}}})$-hash property, where $(\alpha_{\widehat{\mathsf{A}}}, \beta_{\widehat{\mathsf{A}}})$ is defined as

$$\alpha_{\widehat{\mathsf{A}}} \equiv \alpha_{\mathsf{A}} \alpha_{\mathsf{A}'}$$

$$\beta_{\widehat{\mathsf{A}}} \equiv \beta_{\mathsf{A}} + \beta_{\mathsf{A}'}.$$

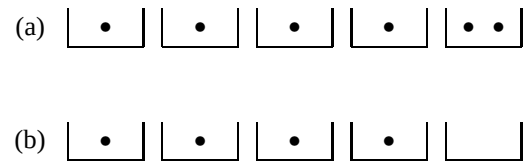

Fig. 4. Properties connecting the number of bins and items (black dots, messages). (a) Saturation property: every bin contains at least one item. (b) Collision-resistance property: every bin contains at most one item.

B. Two Implications of Hash Property

We review two implications of the hash property, which is introduced in [19]. These two implications connect the number of bins and messages (items) and are derived from the hash property by adjusting the number of bins taking account of the number of sequences.

In the following, let $\mathcal{A}$ be a set of functions $A : \mathcal{U}^n \to \mathrm{Im}\mathcal{A}$, where an item is a member of $\mathcal{U}^n$. A function A assigns a label $A\boldsymbol{u}$ of bin to an item $\boldsymbol{u} \in \mathcal{U}^n$.

Saturation property: We prepare a method that finds a typical sequence for each bin. The saturation property is a characteristic of the hash property. Figure 4 (a) represents the ideal situation of this property. When the number of bins is smaller than the number of black dots, we can find a suitable function whereby every bin has at least one black dot. This is because the hash property tends to avoid collisions. It should be noted that it is sufficient for coding problems to satisfy this property for 'almost all (close to probability one)' bins by letting the ratio [the number of bins]/[the number of black dots] be close to zero. To find a typical sequence from each bin, we use the minimum-divergence operation introduced in the construction of codes, where this operation finds a typical sequence when there is. In this situation, the black dots correspond to typical sequences.

We have the following lemma, which is related to the saturation property.

Lemma 2 ([19, Lemma 2]): Assume that the distribution of a random variable $\mathbf{a}$ is uniform on $\mathrm{Im}\mathcal{A}$ and $\mathbf{a}$ and A are mutually independent. If $(\mathcal{A}, p_{\mathsf{A}})$ satisfies (H3'), then

$$p_{\mathsf{A}\mathbf{a}}\left(\{(A, \boldsymbol{a}) : \mathcal{T} \cap \mathcal{C}_A(\boldsymbol{a}) = \emptyset\}\right) \le \alpha_{\mathsf{A}} - 1 + \frac{|\mathrm{Im}\mathcal{A}|\,[\beta_{\mathsf{A}} + 1]}{|\mathcal{T}|}$$

for any $\mathcal{T} \subset \mathcal{U}^n$.

We prove the saturation property from Lemma 2. We have

$$\begin{aligned} E_{\mathsf{A}}\left[p_{\mathbf{c}}\left(\{\boldsymbol{c} : \mathcal{T} \cap \mathcal{C}_{\mathsf{A}}(\boldsymbol{c}) = \emptyset\}\right)\right] &= p_{\mathsf{A}\mathbf{c}}\left(\{(A, \boldsymbol{c}) : \mathcal{T} \cap \mathcal{C}_A(\boldsymbol{c}) = \emptyset\}\right) \\ &\le \alpha_{\mathsf{A}} - 1 + \frac{|\mathrm{Im}\mathcal{A}|\,[\beta_{\mathsf{A}} + 1]}{|\mathcal{T}|}. \end{aligned} \tag{23}$$

By assuming that $|\mathrm{Im}\mathcal{A}|/|\mathcal{T}|$ vanishes as $n \to \infty$, we have the fact that there is a function A such that

$$p_{\mathbf{c}}\left(\{\boldsymbol{c} : \mathcal{T} \cap \mathcal{C}_A(\boldsymbol{c}) = \emptyset\}\right) < \delta$$

for any $\delta > 0$ and sufficiently large n. Since the relation $\mathcal{T} \cap \mathcal{C}_A(\boldsymbol{c}) = \emptyset$ corresponds to an event where there is no $\boldsymbol{u} \in \mathcal{T}$ in bin $\mathcal{C}_A(\boldsymbol{c})$, we have the fact that we can find a member of $\mathcal{T}$ in a randomly selected bin with probability close to one.

Collision-resistance property: A good code assigns a message to a codeword that is different from the codewords of other messages, where the error probability is as small as possible. The collision-resistance property is another characteristic of the hash property. Figure 4 (b) shows the ideal situation as regards this property, where the black dots represent messages we want to distinguish. When the number of bins is greater than the number of black dots, we can find a good function that allocates the black dots to the different bins. This is because the hash property tends to avoid the collision. It should be noted that it is sufficient for coding problems to satisfy this property for 'almost all (close to probability one)' black dots by letting the ratio [the number of black dots]/[the number of bins] be close to zero. This property is used to estimate the decoding error probability. In this situation, the black dots correspond to typical sequences.

We have the following lemma, which is related to to the collision-resistance property.

Lemma 3 ([19, Lemma 1]): If $(\mathcal{A}, p_{\mathsf{A}})$ satisfies (H3'), then

$$p_{\mathsf{A}}\left(\{A : [\mathcal{G} \setminus \{\boldsymbol{u}\}] \cap \mathcal{C}_A(A\boldsymbol{u}) \neq \emptyset\}\right) \leq \frac{|\mathcal{G}|\alpha_{\mathsf{A}}}{|\mathrm{Im}\mathcal{A}|} + \beta_{\mathsf{A}}.$$

for all $\mathcal{G} \subset \mathcal{U}^n$ and $\boldsymbol{u} \in \mathcal{U}^n$.

We prove the collision-resistance property from Lemma 3. Let μ_U be the probability distribution on $\mathcal{G} \subset \mathcal{U}^n$. We have

$$\begin{aligned} E_{\mathsf{A}}\left[\mu_U\left(\{\boldsymbol{u} : [\mathcal{G} \setminus \{\boldsymbol{u}\}] \cap \mathcal{C}_{\mathsf{A}}(\mathsf{A}\boldsymbol{u}) \neq \emptyset\}\right)\right] &\leq \sum_{\boldsymbol{u} \in \mathcal{G}} \mu_U(\boldsymbol{u}) p_{\mathsf{A}}\left(\{A : [\mathcal{G} \setminus \{\boldsymbol{u}\}] \cap \mathcal{C}_A(A\boldsymbol{u}) \neq \emptyset\}\right) \\ &\leq \sum_{\boldsymbol{u} \in \mathcal{G}} \mu_U(\boldsymbol{u}) \left[\frac{|\mathcal{G}|\alpha_{\mathsf{A}}}{|\mathrm{Im}\mathcal{A}|} + \beta_{\mathsf{A}}\right] \\ &\leq \frac{|\mathcal{G}|\alpha_{\mathsf{A}}}{|\mathrm{Im}\mathcal{A}|} + \beta_A. \end{aligned} \tag{24}$$

By assuming that $|\mathcal{G}|/|\mathrm{Im}\mathcal{A}|$ vanishes as $n \to \infty$, we have the fact that there is a function A such that

$$\mu_U\left(\{\boldsymbol{u} : [\mathcal{G} \setminus \{\boldsymbol{u}\}] \cap \mathcal{C}_A(A\boldsymbol{u}) \neq \emptyset\}\right) < \delta$$

for any $\delta > 0$ and sufficiently large n. Since the relation $[\mathcal{G} \setminus \{\boldsymbol{u}\}] \cap \mathcal{C}_A(A\boldsymbol{u}) \neq \emptyset$ corresponds to an event where there is $\boldsymbol{u}' \in \mathcal{G}$ such that $\boldsymbol{u}$ and $\boldsymbol{u}'$ are different members of the same bin (they have the same codeword determined by A), we have the fact that the members of $\mathcal{G}$ are located in different bins (the members of $\mathcal{G}$ can be decoded correctly) with probability close to one.

C. Channel Coding Based on Hash Property

Now, we explain an intuitive construction of a channel code in terms of the saturation property and the collision-resistance property, where the construction is introduced in [19].

We prepare two functions $A : \mathcal{X}^n \to \mathrm{Im}\mathcal{A}$, $B : \mathcal{X}^n \to \mathrm{Im}\mathcal{B}$, and a vector $\boldsymbol{c} \in \mathrm{Im}\mathcal{A}$, and assume that they are shared by an encoder and a decoder. It should be noted that $|\mathrm{Im}\mathcal{A}|$ (resp. $|\mathrm{Im}\mathcal{B}|$) is the number of bins specified by A (resp. B). The function A is analogous to a parity check matrix in the context of linear codes. The set $\mathrm{Im}\mathcal{B}$ is the set of all messages and $|\mathrm{Im}\mathcal{B}|$ is equal to the number of messages.

The flow of vectors is is illustrated in Fig. 5. Let $\boldsymbol{m} \in \mathrm{Im}\mathcal{B}$ be a message, $\boldsymbol{x} \in \mathcal{X}^n$ be a channel input, and $\boldsymbol{y} \in \mathcal{Y}^n$ be a channel output. For $\boldsymbol{c}$ and a message $\boldsymbol{m}$, a function $\widehat{g}_{AB}$ generates a typical sequence $\boldsymbol{x} \in \mathcal{T}_{X,\gamma}$ as a channel input, where $A\boldsymbol{x} = \boldsymbol{c}$ and $B\boldsymbol{x} = \boldsymbol{m}$ are satisfied. The decoder reproduces the channel input $\boldsymbol{x}$ by

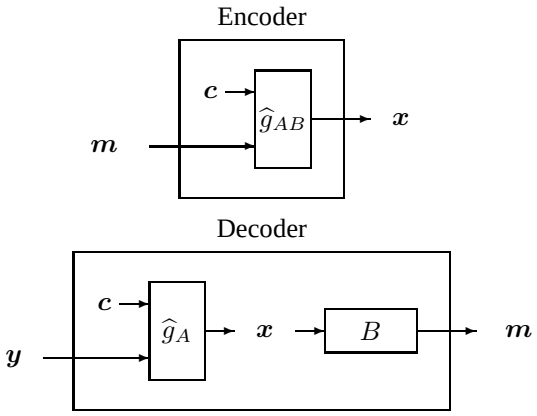

Fig. 5. Construction of Channel Code

using $\widehat{g}_A$ from $\boldsymbol{c}$ and a channel output $\boldsymbol{y}$. Since $(\boldsymbol{x}, \boldsymbol{y})$ is jointly typical and $B\boldsymbol{x} = \boldsymbol{m}$, the decoding succeeds if the amount of information of $\boldsymbol{c}$ is greater than $H(X|Y)$ to satisfy the collision-resistance property. In fact, there are about $2^{nH(X|Y)}$ conditional typical sequences $\boldsymbol{x}$ for given $\boldsymbol{y}$ and it is sufficient to prepare bins specified by A more than $2^{nH(X|Y)}$ to distinguish conditional typical sequences. Formally, this condition corresponds to

$$\frac{\log |\mathrm{Im}\mathcal{A}|}{n} > H(X|Y).$$

On the other hand, the total rate of $\boldsymbol{c}$ and $\boldsymbol{m}$ should be less than $H(X)$ to satisfy the saturating property. Since there is at most $2^{nH(X)}$ typical sequences, it is sufficient to prepare bins specified by A and B less than $2^{nH(X)}$. Formally, this condition corresponds to

$$\frac{\log |\mathrm{Im}\mathcal{A}||\mathrm{Im}\mathcal{B}|}{n} < H(X).$$

Since it is sufficient to satisfy these two inequalities, we have the fact that there is a code when

$$\frac{\log |\mathrm{Im}\mathcal{B}|}{n} < H(X) - H(X|Y).$$

This implies that we can set the encoding rate of messages close to $H(X) - H(X|Y) = I(X;Y)$.

In this paper, we extend this approach to construct a multiple access channel code.

D. Multiple Extension of Collision Resistance Property

To prove the achievability of a multiple access channel code based on the hash property, we extend the lemma related to the collision-resistance property. The hash property is needed to prove the following lemma, and this is another reason why the hash property is introduced. We use the following notations:

$$\mathcal{C}_{A_\mathcal{K}}(\boldsymbol{a}_\mathcal{K}) \equiv \{\boldsymbol{u}_\mathcal{K} : A_j \boldsymbol{u}_j = \boldsymbol{a}_j \text{ for all } j \in \mathcal{K}\}.$$

$$A_\mathcal{K} \boldsymbol{u}_\mathcal{K} \equiv \{A_j \boldsymbol{u}_j\}_{j \in \mathcal{K}}.$$

For $\mathcal{G} \subset \mathcal{U}^n \times \mathcal{V}^n$ and $\boldsymbol{u} \in \mathcal{U}^n$, let $\mathcal{G}_\mathcal{U}$ and $\mathcal{G}_{\mathcal{V}|\mathcal{U}}(\boldsymbol{u})$ be defined as

$$\mathcal{G}_\mathcal{U} \equiv \{\boldsymbol{u} : (\boldsymbol{u}, \boldsymbol{v}) \in \mathcal{G} \text{ for some } \boldsymbol{v} \in \mathcal{V}^n\}$$

$$\mathcal{G}_{\mathcal{V}|\mathcal{U}}(\boldsymbol{u}) \equiv \{\boldsymbol{v} : (\boldsymbol{u}, \boldsymbol{v}) \in \mathcal{G}\}.$$

Furthermore, to shorten the description of the following lemma, we use the following abbreviation

$$\left|\mathcal{G}_{\mathcal{J}|\mathcal{J}^c}\right| \equiv \begin{cases} |\mathcal{G}|, & \text{if } \mathcal{J} = \mathcal{K} \\ \max\limits_{\boldsymbol{u}_{\mathcal{J}^c} \in \mathcal{G}_{\mathcal{U}_{\mathcal{J}^c}}} \left|\mathcal{G}_{\mathcal{U}_{\mathcal{J}}|\mathcal{U}_{\mathcal{J}^c}}\left(\boldsymbol{u}_{\mathcal{J}^c}\right)\right| & \text{otherwise.} \end{cases} \tag{25}$$

for $\mathcal{G} \subset [\mathcal{U}_{\mathcal{K}}]^n$ and $\mathcal{J} \subset \mathcal{K}$. It should be noted that the expression $\left|\mathcal{G}_{\mathcal{J}|\mathcal{J}^c}\right|$ does not represent the cardinality of the set $\mathcal{G}_{\mathcal{J}|\mathcal{J}^c}$.

Lemma 4 ([21, Lemma 7]): For each $j \in \mathcal{K}$, let $\mathcal{A}_j$ be a set of functions $A_j : \mathcal{U}_j^n \to \mathrm{Im}\mathcal{A}_j$ and p_{A_j} be the probability distribution on $\mathcal{A}_j$, where $(\mathcal{A}_j, p_{\mathsf{A}_j})$ satisfies (H3). We assume that random variables $\mathsf{A}_{\mathcal{K}} \equiv \{\mathsf{A}_j\}_{j\in\mathcal{K}}$ are mutually independent. For each $\mathcal{J} \subset \mathcal{K}$, let $\boldsymbol{\alpha}_{\mathsf{A}_{\mathcal{J}}}$ and $\boldsymbol{\beta}_{\mathsf{A}_{\mathcal{J}}}$ be defined as

$$\alpha_{\mathsf{A}_{\mathcal{J}}} \equiv \prod_{j\in\mathcal{J}} \alpha_{\mathsf{A}_j}$$

$$\beta_{\mathsf{A}_{\mathcal{J}}} \equiv \prod_{j\in\mathcal{J}} \left[1 + \beta_{\mathsf{A}_j}\right] - 1.$$

Then

$$p_{\mathsf{A}_{\mathcal{K}}}\left(\{A_{\mathcal{K}} : [\mathcal{G} \setminus \{\boldsymbol{u}_{\mathcal{K}}\}] \cap \mathcal{C}_{A_{\mathcal{K}}}(A_{\mathcal{K}}\boldsymbol{u}_{\mathcal{K}}) \neq \emptyset\}\right) \leq \sum_{\substack{\mathcal{J}\subset\mathcal{K}:\\ \mathcal{J}\neq\emptyset}} \frac{\left|\mathcal{G}_{\mathcal{J}|\mathcal{J}^c}\right| \alpha_{\mathsf{A}_{\mathcal{J}}}\left[\beta_{\mathsf{A}_{\mathcal{J}^c}} + 1\right]}{\prod_{j\in\mathcal{J}} |\mathrm{Im}\mathcal{A}_j|} + \beta_{\mathsf{A}_{\mathcal{K}}}$$

for all $\mathcal{G} \subset [\mathcal{U}_{\mathcal{K}}]^n$ and $\boldsymbol{u}_{\mathcal{K}} \in [\mathcal{U}_{\mathcal{K}}]^n$. Furthermore, if $(\boldsymbol{\alpha}_{\mathsf{A}_j}, \boldsymbol{\beta}_{\mathsf{A}_j})$ satisfies (H1) and (H2) for all $j \in \mathcal{K}$, then

$$\lim_{n\to\infty} \alpha_{\mathsf{A}_{\mathcal{J}}}(n) = 1 \tag{26}$$

$$\lim_{n\to\infty} \beta_{\mathsf{A}_{\mathcal{J}}}(n) = 0 \tag{27}$$

for every $\mathcal{J} \subset \mathcal{K}$.

V. Construction of Codes

In this section, we construct codes for the scenarios introduced in Section III.

A. Private Messages

In this section, we consider a scenario in which k senders transmit independent messages to a receiver and there is no common message to be sent (Fig.1).

First, we construct a code based on the coded time-sharing technique introduced in [13]. For a given $\mu_{Y|X_{\mathcal{K}}}$, μ_U, and $\{\mu_{X_j|U}\}_{j\in\mathcal{K}}$, assume that $R_{\mathcal{K}}$ satisfies (10). Then there is $\{\varepsilon_j\}_{j\in\mathcal{K}}$ such that

$$\sum_{j\in\mathcal{J}} [R_j + \varepsilon_j] < I(X_{\mathcal{J}}; Y|U, X_{\mathcal{J}^c}) - \varepsilon \quad \text{for all } \mathcal{J} \subset \mathcal{K}, \tag{28}$$

where ε is defined as

$$\varepsilon \equiv \eta_{\mathcal{X}_{\mathcal{K}}|\mathcal{U}\mathcal{Y}}\left(2\sum_{j\in\mathcal{K}} \varepsilon_j \middle| 2\sum_{j\in\mathcal{K}} \varepsilon_j\right), \tag{29}$$

where $\eta_{\mathcal{X}_{\mathcal{K}}|\mathcal{U}\mathcal{Y}}$ is defined by (7). For each $j \in \mathcal{K}$, let r_j be defined as

$$r_j \equiv H(X_j|U) - R_j - \varepsilon_j. \tag{30}$$

From (11), (28), and (30), we have

$$r_j \geq I(X_j;Y|U,X_{\mathcal{K}\setminus\{j\}}) - R_j - \varepsilon_j > 0.$$

Let $(\boldsymbol{\mathcal{A}}_j, \boldsymbol{p}_{\mathsf{A}_j})$ and $(\boldsymbol{\mathcal{A}}'_j, \boldsymbol{p}_{\mathsf{A}'_j})$ be ensembles of functions, and let $\mathcal{A}_j \in \boldsymbol{\mathcal{A}}_j$ and $\mathcal{A}'_j \in \boldsymbol{\mathcal{A}}'_j$. Let $A_j \in \mathcal{A}_j$ and $A'_j \in \mathcal{A}'_j$ be functions

$$A_j : \mathcal{X}_j^n \to \mathrm{Im}\mathcal{A}_j$$
$$A'_j : \mathcal{X}_j^n \to \mathrm{Im}\mathcal{A}'_j,$$

respectively. We assume that ensembles satisfy

$$r_j = \frac{\log|\mathrm{Im}\mathcal{A}_j|}{n} \tag{31}$$
$$R_j = \frac{\log|\mathrm{Im}\mathcal{A}'_j|}{n}. \tag{32}$$

For each $j \in \mathcal{K}$, let $\mathcal{M}_j$ be the set of messages defined as

$$\mathcal{M}_j \equiv \mathrm{Im}\mathcal{A}'_j.$$

Then R_j represents the encoding rate of the j-th message. We assume that the j-th encoder and a decoder share functions $A_j \in \mathcal{A}_j$, $A'_j \in \mathcal{A}'_j$ and vectors $\boldsymbol{a}_j \in \mathrm{Im}\mathcal{A}_j$ and $\boldsymbol{u} \in \mathcal{U}^n$.

For each $j \in \mathcal{K}$, we define the j-th encoder as

$$\varphi_j(\boldsymbol{m}_j) \equiv \widehat{g}_{A_jA'_j}(\boldsymbol{a}_j, \boldsymbol{m}_j|\boldsymbol{u})$$

for a message $\boldsymbol{m}_j \in \mathcal{M}_j$, where $\mathcal{C}_{A_jA'_j}(\boldsymbol{a}_j, \boldsymbol{m}_j)$ is defined as

$$\mathcal{C}_{A_jA'_j}(\boldsymbol{a}_j, \boldsymbol{m}_j) \equiv \{\boldsymbol{x}_j : A_j\boldsymbol{x}_j = \boldsymbol{a}_j \text{ and } A'_j\boldsymbol{x}_j = \boldsymbol{m}_j\}. \tag{33}$$

and

$$\widehat{g}_{A_jA'_j}(\boldsymbol{a}_j, \boldsymbol{m}_j|\boldsymbol{u}) \equiv \arg \min_{\boldsymbol{x}'_j \in \mathcal{C}_{A_jA'_j}(\boldsymbol{a}_j,\boldsymbol{m}_j)} D(\nu_{\boldsymbol{x}'_j|\boldsymbol{u}}\|\mu_{X_j|U}|\nu_{\boldsymbol{u}}).$$

We define the decoder as

$$\psi(\boldsymbol{y}) \equiv A'_{\mathcal{K}}\widehat{g}_{A_{\mathcal{K}}}(\boldsymbol{a}_{\mathcal{K}}|\boldsymbol{y},\boldsymbol{u})$$

for a channel output $\boldsymbol{y} \in \mathcal{Y}^n$, where

$$\widehat{g}_{A_{\mathcal{K}}}(\boldsymbol{a}_{\mathcal{K}}|\boldsymbol{y},\boldsymbol{u}) \equiv \arg \min_{\substack{\boldsymbol{x}'_{\mathcal{K}}: \\ \boldsymbol{x}'_j \in \mathcal{C}_{A_j}(\boldsymbol{a}_j) \\ \text{for all } j\in\mathcal{K}}} D(\nu_{\boldsymbol{u}\boldsymbol{x}'_{\mathcal{K}}\boldsymbol{y}}\|\mu_{UX_{\mathcal{K}}Y}).$$

Figure 6 illustrates the code construction for $k = 2$. For given vectors $\boldsymbol{a}_j$, $\boldsymbol{u}$, and a message $\boldsymbol{m}_j$, the function $\widehat{g}_{A_jA'_j}$ finds a conditionally typical sequence $\boldsymbol{x}_j$ satisfying $A_j\boldsymbol{x}_j = \boldsymbol{a}_j$ and $A'_j\boldsymbol{x}_j = \boldsymbol{m}_j$. The function A_j is analogous to the parity check matrix for the j-th message, and the function $\widehat{g}_{A_{\mathcal{K}}}$ is a typical set decoder that guesses the channel input $\boldsymbol{x}_{\mathcal{K}}$ satisfying $A_j\boldsymbol{x}_j = \boldsymbol{a}_j$ for all $j \in \mathcal{K}$, where vectors $\boldsymbol{a}_{\mathcal{K}}$, $\boldsymbol{u}$, and a channel output $\boldsymbol{y}$ are given.

Here, let us remark on the relations (28) and (30). From these relations and (11), we have

$$r_j + R_j = H(X_j|U) - \varepsilon_j \quad \text{for all } j \in \mathcal{K} \tag{34}$$

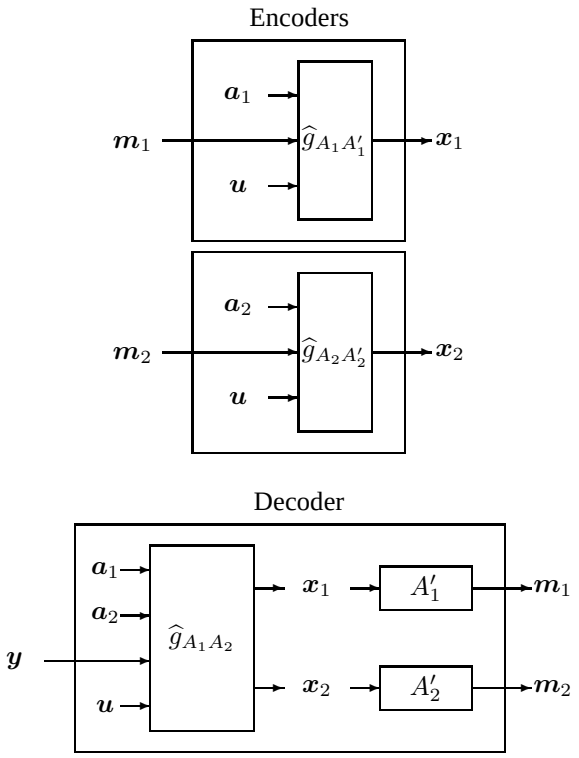

Fig. 6. Construction of Multiple Access Channel Code: Private Messages (Coded Time-sharing)

and

$$\begin{aligned}\sum_{j\in\mathcal{J}} r_j &= \sum_{j\in\mathcal{J}} [H(X_j|U) - R_j - \varepsilon_j] \\ &= H(X_{\mathcal{J}}|U) - \sum_{j\in\mathcal{J}} [R_j + \varepsilon_j] \\ &> H(X_{\mathcal{J}}|U) - I(X_{\mathcal{J}}; Y|U, X_{\mathcal{J}^c}) + \varepsilon \\ &= H(X_{\mathcal{J}}|U, X_{\mathcal{J}^c}, Y) + \varepsilon \end{aligned} \tag{35}$$

for all $\mathcal{J} \subset \mathcal{K}$. Condition (34) is sufficient for the saturation property, that is, for a given $\boldsymbol{u}$ the j-th encoder can find a conditionally typical sequence corresponding to the j-th message $\boldsymbol{m}_j$ when the number $2^{n[r_j+R_j]}$ of bins is smaller than the number of typical sequences. Condition (35) is sufficient for the collision-resistance property, that is, the decoding error probability goes to zero if the rate vector $r_{\widetilde{\mathcal{K}}}$ of the vector $\boldsymbol{a}_{\mathcal{K}}$ is in the Slepian-Wolf region of the correlated source coding. It should be noted that the decoder can recover messages $\boldsymbol{m}_{\mathcal{K}}$ when the channel input $\boldsymbol{x}_{\mathcal{K}}$ is successfully decoded by operating $A'_{\mathcal{K}}$ to $\boldsymbol{x}_{\mathcal{K}}$ because the j-th message $\boldsymbol{m}_j$ satisfies $A'_j \boldsymbol{x}_j = \boldsymbol{m}_j$.

For each $j \in \mathcal{K}$, let M_j be a random variable corresponding to the j-th message, where the probability distribution p_{M_j} is uniform on $\mathcal{M}_j$. Let $\mathrm{Error}(A_{\mathcal{K}}, A'_{\mathcal{K}}, \boldsymbol{a}_{\mathcal{K}})$ be the decoding error probability. We have the following theorem.

Theorem 1: Let $\mu_{Y|X_{\mathcal{K}}}$ be the conditional probability distribution of a stationary memoryless channel and

$\mu_{UX_{\mathcal{K}}Y}$ be defined by (11) for given probability distributions μ_U and $\{\mu_{X_j|U}\}_{j\in\mathcal{K}}$. For given $R_{\mathcal{K}} \in \mathcal{R}(\mu_U, \{\mu_{X_j|U}\}_{j\in\mathcal{K}})$ and $\{\varepsilon_j\}_{j\in\mathcal{K}}$ satisfying (31)–(29), assume that ensembles $(\boldsymbol{\mathcal{A}}_j, \boldsymbol{p}_{\mathsf{A}_j})$ and $(\boldsymbol{\mathcal{A}}'_j, \boldsymbol{p}_{\mathsf{A}'_j})$ have a hash property for all $j \in \mathcal{K}$. Then, for any $\delta > 0$ and all sufficiently large n, there are functions (sparse matrices) $\{A_j\}_{j\in\mathcal{K}}$, $\{A'_j\}_{j\in\mathcal{K}}$, and vectors $\{\boldsymbol{a}_j\}_{j\in\mathcal{K}}$, $\boldsymbol{u}$ such that $A_j \in \mathcal{A}_j$, $A'_j \in \mathcal{A}'_j$, $\boldsymbol{a}_j \in \mathrm{Im}\mathcal{A}_j$, $\boldsymbol{u} \in \mathcal{U}^n$, and

$$\mathrm{Error}(A_{\mathcal{K}}, A'_{\mathcal{K}}, \boldsymbol{a}_{\mathcal{K}}, \boldsymbol{u}) < \delta. \tag{36}$$

Next, we construct a code with $R_{\mathcal{K}} \in \mathcal{R}(\{\mu_{X_j}\}_{j\in\mathcal{K}})$ by letting U be a constant, that is, $|\mathcal{U}| = 1$. Although the result is straightforward, we describe the corollary which is used in the next section. Condition (10) is replaced by (8). Condition (28) is replaced by

$$\sum_{j\in\mathcal{J}}[R_j + \varepsilon_j] < I(X_{\mathcal{J}}; Y|X_{\mathcal{J}^c}) - \varepsilon \quad \text{for all } \mathcal{J} \subset \mathcal{K}, \tag{37}$$

where ε is defined as

$$\varepsilon \equiv \eta_{\mathcal{X}_{\mathcal{K}}|\mathcal{Y}}\left(2\sum_{j\in\mathcal{K}}\varepsilon_j \middle| 2\sum_{j\in\mathcal{K}}\varepsilon_j\right). \tag{38}$$

Definition (30) is replaced by

$$r_j = H(X_j) - R_j - \varepsilon_j \quad \text{for all } j \in \mathcal{K}. \tag{39}$$

Functions $\widehat{g}_{A_jA'_j}$ and $\widehat{g}_{A_{\mathcal{K}}}$ can be replaced by

$$\widehat{g}_{A_jA'_j}(\boldsymbol{a}_j, \boldsymbol{m}_j) \equiv \arg\min_{\boldsymbol{x}'_j\in\mathcal{C}_{A_jA'_j}(\boldsymbol{a}_j,\boldsymbol{m}_j)} D(\nu_{\boldsymbol{x}'_j}\|\mu_{X_j})$$

$$\widehat{g}_{A_{\mathcal{K}}}(\boldsymbol{a}_{\mathcal{K}}|\boldsymbol{y}) \equiv \arg \min_{\substack{\boldsymbol{x}'_{\mathcal{K}}:\\ \boldsymbol{x}'_j\in\mathcal{C}_{A_j}(\boldsymbol{a}_j)\\ \text{for all } j\in\mathcal{K}}} D(\nu_{\boldsymbol{x}'_{\mathcal{K}}\boldsymbol{y}}\|\mu_{X_{\mathcal{K}}Y}),$$

respectively, where $\mathcal{C}_{A_jA'_j}(\boldsymbol{a}_j, \boldsymbol{m}_j)$ is defined by (33).

Figure 7 illustrates the code construction for $k = 2$. We have the following corollary.

Corollary 2: Let $\mu_{Y|X_{\mathcal{K}}}$ be the conditional probability distribution of a stationary memoryless channel and $\mu_{X_{\mathcal{K}}Y}$ be defined by (9) for a given $\{\mu_{X_j}\}_{j\in\mathcal{K}}$. For given $R_{\mathcal{K}} \in \mathcal{R}(\{\mu_{X_j}\}_{j\in\mathcal{K}})$ and $\{\varepsilon_j\}_{j\in\mathcal{K}}$, satisfying (31), (32), and (39)–(38), assume that ensembles $(\boldsymbol{\mathcal{A}}_j, \boldsymbol{p}_{\mathsf{A}_j})$ and $(\boldsymbol{\mathcal{A}}'_j, \boldsymbol{p}_{\mathsf{A}'_j})$ have a hash property for all $j \in \mathcal{K}$. Then, for any $\delta > 0$ and all sufficiently large n, there are functions $\{A_j\}_{j\in\mathcal{K}}$, $\{A'_j\}_{j\in\mathcal{K}}$, and vectors $\{\boldsymbol{a}_j\}_{j\in\mathcal{K}}$ such that $A_j \in \mathcal{A}_j$, $A'_j \in \mathcal{A}'_j$, $\boldsymbol{a}_j \in \mathrm{Im}\mathcal{A}_j$, and $\mathrm{Error}(A_{\mathcal{K}}, A'_{\mathcal{K}}, \boldsymbol{a}_{\mathcal{K}}) < \delta$, where $\mathrm{Error}(A_{\mathcal{K}}, A'_{\mathcal{K}}, \boldsymbol{a}_{\mathcal{K}})$ denotes the error probability.

B. Multiple Common Messages

In the following, we consider the scenario (Fig.2) where there are $\widetilde{k}$ messages and k senders transmit messages common to some users.

In the following, we assume that for given $\mu_{Y|X_{\mathcal{K}}}$, $\{\mu_{\widetilde{X}_i}\}_{i\in\widetilde{\mathcal{K}}}$, and $\{f_j\}_{j\in\mathcal{K}}$, the rate vector $R_{\widetilde{\mathcal{K}}}$ satisfies $R_{\widetilde{\mathcal{K}}} \in \mathcal{R}_{\mathrm{H}}(\{\mu_{\widetilde{X}_i}\}_{i\in\widetilde{\mathcal{K}}}, \{f_j\}_{j\in\mathcal{K}})$. For a given k-input multiple access channel $\mu_{Y|X_{\mathcal{K}}}$, let us consider a $\widetilde{k}$-input multiple access channel $\mu_{Y|\widetilde{X}_{\widetilde{\mathcal{K}}}}$ defined as

$$\mu_{Y|\widetilde{X}_{\widetilde{\mathcal{K}}}}(y|\widetilde{x}_{\widetilde{\mathcal{K}}}) \equiv \sum_{x_{\mathcal{K}}} \mu_{Y|X_{\mathcal{K}}}(y|x_{\mathcal{K}}) \prod_{j\in\mathcal{K}} \chi(f_j(\widetilde{x}_{\widetilde{\mathcal{K}}_j}) = x_j).$$

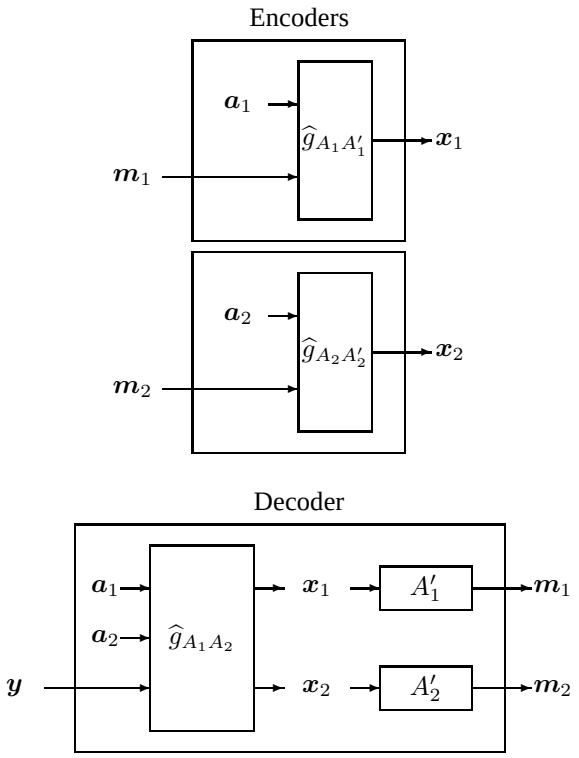

Fig. 7. Construction of Multiple Access Channel Code: Private Messages

Then the scenario of multiple common messages for the channel $\mu_{Y|\mathcal{X}_\mathcal{K}}$ can be reduced to the scenario of private messages for the channel $\mu_{Y|\widetilde{X}_{\widetilde{\mathcal{K}}}}$ in which the i-th input terminal has access to its private message M_i and there is no common message. Then, by applying Corollary 2 to the channel $\mu_{Y|\widetilde{X}_{\widetilde{\mathcal{K}}}}$, we have the fact that there is a code $(\varphi_{\widetilde{\mathcal{K}}}, \psi)$ for this channel at $R_{\widetilde{\mathcal{K}}}$ satisfying (14). Figure 8 illustrates the construction of the code for the channel $\mu_{Y|\widetilde{X}_{\widetilde{\mathcal{K}}}}$. A code $(\widetilde{\varphi}_\mathcal{K}, \widetilde{\psi})$ for the channel $\mu_{Y|X_\mathcal{K}}$ is given as

$$\widetilde{\varphi}_j(\boldsymbol{m}_{\widetilde{\mathcal{K}}_j}) \equiv \boldsymbol{f}_j\left(\{\varphi_i(\boldsymbol{m}_i)\}_{i\in\widetilde{\mathcal{K}}_j}\right)$$
$$\widetilde{\psi}(\boldsymbol{y}) \equiv \psi(\boldsymbol{y})$$

for a multiple message $\boldsymbol{m}_{\widetilde{\mathcal{K}}}$, where

$$\boldsymbol{f}_j(\widetilde{\boldsymbol{x}}_{\widetilde{\mathcal{K}}_j}) \equiv (f_j(\widetilde{x}_{\widetilde{\mathcal{K}}_j,1}), \ldots, f_j(\widetilde{x}_{\widetilde{\mathcal{K}}_j,n}))$$

for each $j \in \mathcal{K}$ and $\widetilde{\boldsymbol{x}}_{\widetilde{\mathcal{K}}_j} \equiv \{\widetilde{\boldsymbol{x}}_i\}_{i\in\widetilde{\mathcal{K}}_j}$. Figure 9 illustrates the construction of the j-th encoder, where we define $\widetilde{k}_j \equiv |\widetilde{\mathcal{K}}_j|$.

C. Two-user Multiple Access Channel Coding: Private and Common Messages

In this section we consider a scenario (Fig.3) in which one of two senders has access to messages M_0 and M_1 and another sender has access to messages M_0 and M_2. We construct a code based on a method that is analogous to a superposition coding.

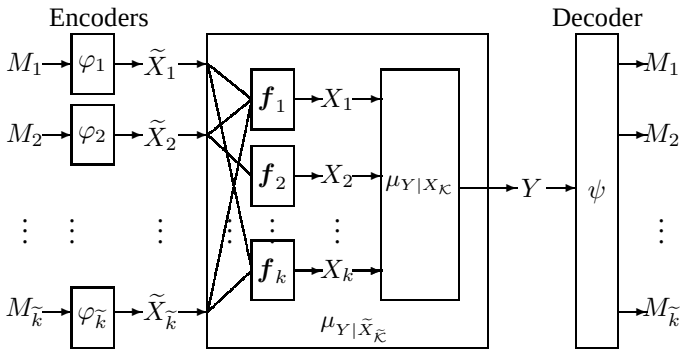

Fig. 8. Reduction of Multiple Common Messages to Private Messages

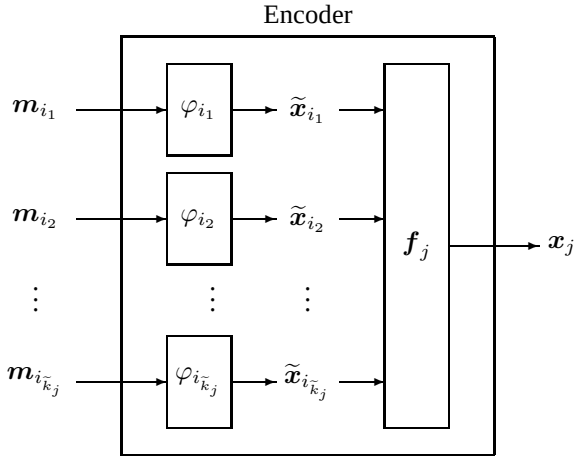

Fig. 9. Construction of j-th Encoder

For given $\mu_{Y|X_1X_2}$, $\mu_{X_1|X_0}$, $\mu_{X_2|X_0}$, and μ_{X_0}, assume that (R_0, R_1, R_2) satisfies (16)–(20) and the following three conditions

$$R_0 < I(X_0; X_1, X_2, Y) \tag{40}$$

$$R_0 + R_1 < I(X_0, X_1; X_2, Y) \tag{41}$$

$$R_0 + R_2 < I(X_0, X_2; X_1, Y), \tag{42}$$

which will be eliminated by using the rate-splitting technique introduced later. Then there is ε such that $\varepsilon > 0$ and

$$R_0 + \varepsilon_0 < I(X_0; X_1, X_2, Y) - \varepsilon \tag{43}$$

$$R_1 + \varepsilon_1 < I(X_1; Y|X_0, X_2) - \varepsilon \tag{44}$$

$$R_2 + \varepsilon_1 < I(X_2; Y|X_0, X_1) - \varepsilon \tag{45}$$

$$R_0 + R_1 + \varepsilon_0 + \varepsilon_1 < I(X_0, X_1; X_2, Y) - \varepsilon \tag{46}$$

$$R_0 + R_2 + \varepsilon_0 + \varepsilon_2 < I(X_0, X_2; X_1, Y) - \varepsilon \tag{47}$$

$$R_1 + R_2 + \varepsilon_1 + \varepsilon_2 < I(X_1, X_2; Y|X_0) - \varepsilon \tag{48}$$

$$R_0 + R_1 + R_2 + \varepsilon_0 + \varepsilon_1 + \varepsilon_2 < I(X_1, X_2; Y) - \varepsilon, \tag{49}$$

where ε is defined by

$$\varepsilon \equiv \eta_{\mathcal{X}_{\widetilde{\mathcal{K}}}|\mathcal{Y}} \left(2\sum_{j\in\widetilde{\mathcal{K}}} \varepsilon_j \middle| 2\sum_{j\in\widetilde{\mathcal{K}}} \varepsilon_j \right). \tag{50}$$

For each $j \in \widetilde{\mathcal{K}} \equiv \{0, 1, 2\}$, let r_j be defined as

$$r_0 \equiv H(X_0) - R_0 - \varepsilon_0 \tag{51}$$

$$r_1 \equiv H(X_1|X_0) - R_1 - \varepsilon_1 \tag{52}$$

$$r_2 \equiv H(X_2|X_0) - R_2 - \varepsilon_2. \tag{53}$$

From (21), (30), and (43)–(45), we have

$$r_0 \geq I(X_0; X_1, X_2, Y) - R_0 - \varepsilon_0 > 0$$

$$r_1 \geq I(X_1; Y|X_0, X_2) - R_1 - \varepsilon_1 > 0$$

$$r_2 \geq I(X_2; Y|X_0, X_1) - R_2 - \varepsilon_2 > 0.$$

For $i \in \widetilde{\mathcal{K}}$, let $(\boldsymbol{\mathcal{A}}_i, \boldsymbol{p}_{\mathsf{A}_i})$ and $(\boldsymbol{\mathcal{A}}'_i, \boldsymbol{p}_{\mathsf{A}'_i})$ be ensembles of functions, and let $\mathcal{A}_i \in \boldsymbol{\mathcal{A}}_i$ and $\mathcal{A}'_i \in \boldsymbol{\mathcal{A}}'_i$. Let $A_i \in \mathcal{A}_i$ and $A'_i \in \mathcal{A}'_i$ be functions

$$A_i : \mathcal{X}_i^n \to \mathrm{Im}\mathcal{A}_i$$

$$A'_i : \mathcal{X}_i^n \to \mathrm{Im}\mathcal{A}'_i,$$

respectively. We assume that ensembles satisfy

$$r_i = \frac{\log|\mathrm{Im}\mathcal{A}_i|}{n} \tag{54}$$

$$R_i = \frac{\log|\mathrm{Im}\mathcal{A}'_i|}{n}. \tag{55}$$

Let $\mathcal{M}_i$ be the set of messages defined as

$$\mathcal{M}_i \equiv \mathrm{Im}\mathcal{A}'_i.$$

Then (R_0, R_1, R_2) represents the encoding rate of this code. We assume that, for each $j \in \mathcal{K} \equiv \{1, 2\}$, the j-th encoder and a decoder share functions $A_0 \in \mathcal{A}_0$, $A'_0 \in \mathcal{A}'_0$, $A_j \in \mathcal{A}_j$, $A'_j \in \mathcal{A}'_j$, and vectors $\boldsymbol{a}_0 \in \mathrm{Im}\mathcal{A}_0$ and $\boldsymbol{a}_j \in \mathrm{Im}\mathcal{A}_j$.

Let $(\boldsymbol{m}_0, \boldsymbol{m}_1, \boldsymbol{m}_2) \in \mathcal{M}_{\widetilde{\mathcal{K}}}$, be a multiple message. For each $j \in \mathcal{K}$, we define the j-th encoder as

$$\varphi_j(\boldsymbol{m}_0, \boldsymbol{m}_j) \equiv g_{A_j A'_j}(\boldsymbol{a}_j, \boldsymbol{m}_j | g_{A_0 A'_0}(\boldsymbol{a}_0, \boldsymbol{m}_0)),$$

where

$$g_{A_0 A'_0}(\boldsymbol{a}_0, \boldsymbol{m}_0) \equiv \arg \min_{\boldsymbol{x}'_0 \in \mathcal{C}_{A_0 A'_0}(\boldsymbol{a}_0, \boldsymbol{m}_0)} D(\nu_{\boldsymbol{x}'_0} \| \mu_{X_0})$$

$$g_{A_j A'_j}(\boldsymbol{a}_j, \boldsymbol{m}_j | \boldsymbol{x}_0) \equiv \arg \min_{\boldsymbol{x}'_j \in \mathcal{C}_{A_j A'_j}(\boldsymbol{a}_j, \boldsymbol{m}_j)} D(\nu_{\boldsymbol{x}'_j|\boldsymbol{x}_0} \| \mu_{X_j|X_0} | \nu_{\boldsymbol{x}_0}),$$

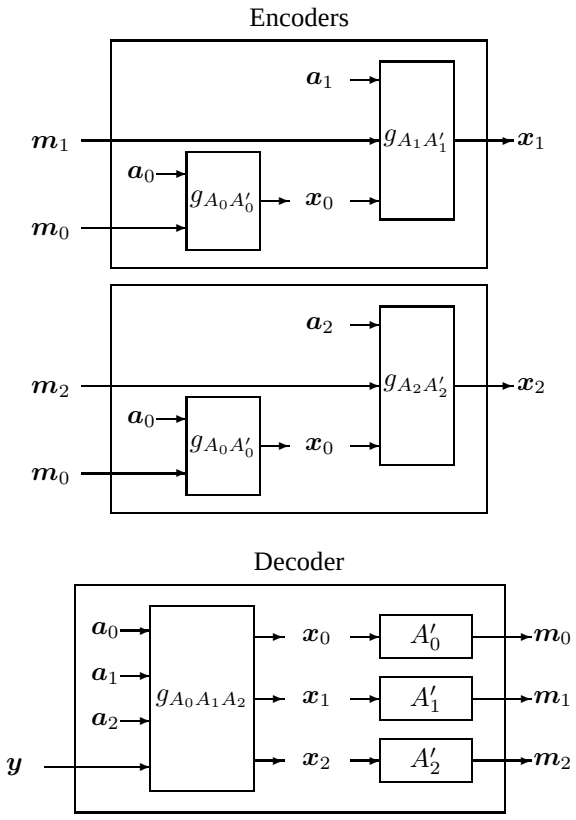

Fig. 10. Construction of Multiple Access Channel Code: Private and Common Messages

where $\mathcal{C}_{A_jA_j'}(\boldsymbol{a}_j, \boldsymbol{m}_j)$ is defined by (33). We define the decoder as

$$\psi(\boldsymbol{y}) \equiv (A_0', A_1', A_2') g_{A_0A_1A_2}(\boldsymbol{a}_0, \boldsymbol{a}_1, \boldsymbol{a}_2 | \boldsymbol{y})$$

for a channel output $\boldsymbol{y} \in \mathcal{Y}^n$, where

$$g_{A_0A_1A_2}(\boldsymbol{a}_0, \boldsymbol{a}_1, \boldsymbol{a}_2 | \boldsymbol{y}) \equiv \arg \min_{\substack{(\boldsymbol{x}_0', \boldsymbol{x}_1', \boldsymbol{x}_2'): \\ \boldsymbol{x}_0' \in \mathcal{C}_{A_0}(\boldsymbol{a}_0) \\ \boldsymbol{x}_1' \in \mathcal{C}_{A_1}(\boldsymbol{a}_1) \\ \boldsymbol{x}_2' \in \mathcal{C}_{A_2}(\boldsymbol{a}_2)}} D(\nu_{\boldsymbol{x}_0'\boldsymbol{x}_1'\boldsymbol{x}_2'\boldsymbol{y}} \| \mu_{X_0X_1X_2Y})$$

$$(A_0', A_1', A_2')(\boldsymbol{x}_0, \boldsymbol{x}_1, \boldsymbol{x}_2) \equiv (A_0'\boldsymbol{x}_0, A_1'\boldsymbol{x}_1, A_2'\boldsymbol{x}_2).$$

Figure 10 illustrates the code construction. It should be noted that the construction is analogous to the superposition coding introduced in [4], where the function $\widehat{g}_{A_0A_0'}$ finds a cloud center $\boldsymbol{x}_0$ and the function $\widehat{g}_{A_jA_j'}$ finds a satellite $\boldsymbol{x}_j$ of the cloud center $\boldsymbol{x}_0$.

Here, we remark on relations (30) and (43)–(49). From these conditions and (21), we have

$$r_0 + R_0 = H(X_0) - \varepsilon_0 \tag{56}$$

$$r_1 + R_1 = H(X_1|X_0) - \varepsilon_1 \tag{57}$$

$$r_2 + R_2 = H(X_2|X_0) - \varepsilon_2 \tag{58}$$

and

$$r_0 > H(X_0|X_1, X_2, Y) + \varepsilon \tag{59}$$

$$r_1 > H(X_1|X_0, X_2, Y) + \varepsilon \tag{60}$$

$$r_2 > H(X_2|X_0, X_1, Y) + \varepsilon \tag{61}$$

$$r_0 + r_1 > H(X_0, X_1|X_2, Y) + \varepsilon \tag{62}$$

$$r_0 + r_2 > H(X_0, X_2|X_1, Y) + \varepsilon \tag{63}$$

$$r_1 + r_2 > H(X_1, X_2|X_0, Y) + \varepsilon \tag{64}$$

$$r_0 + r_1 + r_2 > H(X_0, X_1, X_2|Y) + \varepsilon. \tag{65}$$

Conditions (56)–(58) are sufficient for the saturation property, that is, $\widehat{g}_{A_0 A_0'}$ can find a typical sequence corresponding to the message $\boldsymbol{m}_0$, when the number $2^{n[r_0+R_0]}$ of bins is smaller than the number of typical sequences. Similarly, $\widehat{g}_{A_i A_i'}$ can find a conditionally typical sequence for a given $\boldsymbol{x}_0$ corresponding to the i-th message $\boldsymbol{m}_i$ when the number $2^{n[r_i+R_i]}$ of bins is smaller than the number of conditionally typical sequences. Conditions (59)–(65) are sufficient for the collision-resistance property, that is, the decoding error probability goes to zero if the rate $r_{\mathcal{K}}$ of the vector $\boldsymbol{a}_{\mathcal{K}}$ is in the Slepian-Wolf region of the correlated source coding. It should be noted that when the channel input $(\boldsymbol{x}_0, \boldsymbol{x}_1, \boldsymbol{x}_2)$ is successfully decoded the decoder can recover the i-th message $\boldsymbol{m}_i$ by operating A_i' to $\boldsymbol{x}_i$ because $\boldsymbol{m}_i$ satisfies $A_i' \boldsymbol{x}_i = \boldsymbol{m}_i$.

For each $i \in \widetilde{\mathcal{K}}$, let M_i be a random variable corresponding to the i-th message, where the probability distribution p_{M_i} is uniform on $\mathcal{M}_i$. Let $\mathrm{Error}(A_{\widetilde{\mathcal{K}}}, A'_{\widetilde{\mathcal{K}}}, \boldsymbol{a}_{\widetilde{\mathcal{K}}})$ be the error probability of this code. We have the following theorem.

Theorem 3: Let $\mu_{Y|X_1X_2}$ be the conditional probability distribution of a stationary memoryless channel and $\mu_{X_0X_1X_2Y}$ be defined by (21) for given probability distributions μ_{X_0}, $\mu_{X_1|X_0}$, and $\mu_{X_2|X_0}$. For given (r_0, r_1, r_2), (R_0, R_1, R_2), and $(\varepsilon_0, \varepsilon_1, \varepsilon_2)$ satisfying (30) and (43)–(50), assume that ensembles $(\boldsymbol{\mathcal{A}}_j, \boldsymbol{p}_{\mathsf{A}_j})$ and $(\boldsymbol{\mathcal{A}}'_j, \boldsymbol{p}_{\mathsf{A}'_j})$ have a hash property for all $j \in \widetilde{\mathcal{K}}$. Then, for any $\delta > 0$ and all sufficiently large n, there are functions $\{A_j\}_{j\in\widetilde{\mathcal{K}}}$, $\{A'_j\}_{j\in\widetilde{\mathcal{K}}}$ and vectors $\{\boldsymbol{a}_j\}_{j\in\widetilde{\mathcal{K}}}$ such that $A_j \in \mathcal{A}_j$, $A'_j \in \mathcal{A}'_j$, $\boldsymbol{a}_j \in \mathrm{Im}\mathcal{A}_j$, and

$$\mathrm{Error}(A_{\widetilde{\mathcal{K}}}, A'_{\widetilde{\mathcal{K}}}, \boldsymbol{a}_{\widetilde{\mathcal{K}}}) < \delta. \tag{66}$$

In the following, we employ a rate splitting technique to eliminate conditions (40)–(42). Assume that (R_0, R_1, R_2) satisfies conditions (16)–(20) and (40)–(42). From Theorem 3, we have the fact that there is a code with encoding rate (R_0, R_1, R_2). Let $\boldsymbol{m}_0 \in \mathcal{X}_0^{nR_0}$ be a common message and $\boldsymbol{m}_1 \in \mathcal{X}_1^{nR_1}$ and $\boldsymbol{m}_2 \in \mathcal{X}_2^{nR_2}$ be the private messages of two different encoders. We divide the private messages into two parts

$$\boldsymbol{m}_1 = \left(m_1^{n[R_1-R_1'']}, m_1^{nR_1''}\right)$$
$$\boldsymbol{m}_2 = \left(m_1^{n[R_2-R_2'']}, m_2^{nR_2''}\right),$$

where (R_1'', R_2'') satisfies

$$0 \le R_1'' \le R_1 \tag{67}$$

$$0 \le R_2'' \le R_2. \tag{68}$$

Let us interpret $\left(\boldsymbol{m}_0, m_1^{nR_1''}, m_2^{nR_2''}\right)$ as the common message and $m_j^{n[R_j-R_j'']}$ as the private message of the j-th encoder. Then we have the fact that rate (R_0', R_1', R_2') satisfying

$$R_0' = R_0 + R_1'' + R_2'' \tag{69}$$

$$R_1' = R_1 - R_1'' \tag{70}$$

$$R_2' = R_2 - R_2'' \tag{71}$$

is achievable by using the same code obtained from Theorem 3.

Now we prove the fact that for all $(R_0', R_1', R_2') \in \mathcal{R}_{\mathrm{SW}}(\mu_{X_0}, \mu_{X_1|X_0}, \mu_{X_2|X_0})$ there is a pair (R_1'', R_2'') such that (R_0, R_1, R_2) satisfies conditions (16)–(20), (40)–(42), (67)–(71). From (69)–(71), we have

$$R_0 = R_0' - R_1'' - R_2'' \tag{72}$$

$$R_1 = R_1' + R_1'' \tag{73}$$

$$R_2 = R_2' + R_2''. \tag{74}$$

By substituting these inequalities into (16)–(20), (40)–(42), (67), and (68), we have

$$R_0' - R_1'' - R_2'' \geq 0$$

$$0 \leq R_1' + R_1'' < I(X_1; Y|X_0, X_2)$$

$$0 \leq R_2' + R_2'' < I(X_2; Y|X_0, X_1)$$

$$R_1' + R_1'' + R_2' + R_2'' < I(X_1, X_2; Y|X_0)$$

$$R_0' + R_1' + R_2' < I(X_0, X_1, X_2; Y)$$

$$R_0' - R_1'' - R_2'' < I(X_0; X_1, X_2, Y)$$

$$R_0' + R_1' - R_2'' < I(X_0, X_1; X_2, Y)$$

$$R_0' + R_2' - R_1'' < I(X_0, X_2; X_1, Y)$$

$$0 \leq R_1'' \leq R_1' + R_1''$$

$$0 \leq R_2'' \leq R_2' + R_2'',$$

where we use the relation $I(X_0, X_1, X_2; Y) = I(X_1, X_2; Y)$ obtained from (21) in the fifth inequality. By eliminating R_1'' and R_2'' from these inequalities by using the Fourier-Motzkin method (see [10, Appendix D][28]) and the relation $I(X_1; X_2|X_0) = 0$ obtained from (21), we have the fact that $(R_0', R_1', R_2') \in \mathcal{R}_{\mathrm{SW}}(\mu_{X_0}, \mu_{X_1|X_0}, \mu_{X_2|X_0})$. This implies that for $(R_0', R_1', R_2') \in \mathcal{R}_{\mathrm{SW}}(\mu_{X_0}, \mu_{X_1|X_0}, \mu_{X_2|X_0})$, there is (R_1'', R_2'') such that (R_0, R_1, R_2) defined by (72)–(74) satisfies $(R_0, R_1, R_2) \in \mathcal{R}_{\mathrm{SW}}(\mu_{X_0}, \mu_{X_1|X_0}, \mu_{X_2|X_0})$, (40)–(42), (67), and (68). This means that we can construct codes with $(R_0', R_1', R_2') \in \mathcal{R}_{\mathrm{SW}}(\mu_{X_0}, \mu_{X_1|X_0}, \mu_{X_2|X_0})$. Thus, conditions (40)–(42) are eliminated.

VI. Proof of Theorems

In this section, we prove Theorems 1 and 3. Before describing the proof, we remark on the outline of the proof. The proof is similar to the conventional random coding argument, where a codebook is randomly

generated and it is proved that the average error probability tends to zero as n goes to infinity. However, there is a definite difference from the conventional random coding argument in the following proof because our proof is based on random partitioning and the probability distribution of a codebook is different. This will be explained in detail later.

In the following proof, we omit the dependence of X, Y, U on n when they appear in the subscript of μ. For $\boldsymbol{u} \equiv (u_1, \ldots, u_n) \in \mathcal{U}^n$ and $(\boldsymbol{x}_{\mathcal{K}}, \boldsymbol{y}) \equiv ((\{x_{1,j}\}_{j\in\mathcal{K}}, y_1), \ldots, (\{x_{n,j}\}_{j\in\mathcal{K}}, y_n)) \in [\mathcal{X}_{\mathcal{K}}]^n \times \mathcal{Y}^n$, $\mu_U(\boldsymbol{u})$ and $\mu_{Y|X_{\mathcal{K}}}(\boldsymbol{y}|\boldsymbol{x}_{\mathcal{K}})$ are defined as

$$\mu_U(\boldsymbol{u}) \equiv \prod_{i=1}^{n} \mu_U(u_i)$$
$$\mu_{Y|X_{\mathcal{K}}}(\boldsymbol{y}|\boldsymbol{x}_{\mathcal{K}}) \equiv \prod_{i=1}^{n} \mu_{Y|X_{\mathcal{K}}}(y_i|\{x_{i,j}\}_{j\in\mathcal{K}}).$$

A. Proof of Theorem 1

In the following, we assume that ensembles $(\boldsymbol{\mathcal{A}}_j, \boldsymbol{p}_{\mathsf{A}_j})$ and $(\boldsymbol{\mathcal{A}}'_j, \boldsymbol{p}_{\mathsf{A}'_j})$ have a hash property for all $j \in \mathcal{K}$. Then, from Lemma 1, ensemble $(\widehat{\boldsymbol{\mathcal{A}}}_j, \boldsymbol{p}_{\widehat{\mathsf{A}}_j})$ defined by

$$\widehat{A}_j \boldsymbol{x}_j \equiv (A_j \boldsymbol{x}_j, A'_j \boldsymbol{x}_j)$$

has a $(\boldsymbol{\alpha}_{\widehat{\mathsf{A}}_j}, \boldsymbol{\beta}_{\widehat{\mathsf{A}}_j})$-hash property, where

$$p_{\widehat{A}_j}(\widehat{A}_j) \equiv p_{A_j}(A_j) p_{A'_j}(A'_j)$$
$$\alpha_{\widehat{\mathsf{A}}_j} \equiv \alpha_{\mathsf{A}_j} \alpha_{\mathsf{A}'_j}$$
$$\beta_{\widehat{\mathsf{A}}_j} \equiv \beta_{\mathsf{A}_j} + \beta_{\mathsf{A}'_j}.$$

Since

$$\lim_{n\to\infty} \beta_{\mathsf{A}_{\mathcal{K}}}(n) = \lim_{n\to\infty} \left[\prod_{j\in\mathcal{K}} [\beta_{\mathsf{A}_j} + 1] - 1 \right]$$
$$= 0, \tag{75}$$

there is a sequence $\boldsymbol{\kappa} \equiv \{\kappa(n)\}_{n=1}^{\infty}$ such that

$$\lim_{n\to\infty} \kappa(n) = \infty \tag{76}$$
$$\lim_{n\to\infty} [\kappa(n)]^k \beta_{\mathsf{A}_{\mathcal{K}}}(n) = 0 \tag{77}$$
$$\lim_{n\to\infty} \frac{\log \kappa(n)}{n} = 0, \tag{78}$$

where k is the number of encoders. For example, we obtain such a $\boldsymbol{\kappa}$ by letting

$$\kappa(n) \equiv \begin{cases} n^{\xi/k} & \text{if } \exists \xi > 0 \text{ s.t. } \beta_{\mathsf{A}_{\mathcal{K}}}(n) = o\left(n^{-\xi/k}\right) \\ [\beta_{\mathsf{A}_{\mathcal{K}}}(n)]^{-1/[k+1]} & \text{otherwise} \end{cases}$$

for every n. If $\beta_{\mathsf{A}_{\mathcal{K}}}(n)$ is not $o\left(n^{-\xi/k}\right)$, there is a κ' such that $\kappa' > 0$, $\beta_{\mathsf{A}_{\mathcal{K}}}(n) n^{\xi/k} > \kappa'$ and

$$\frac{\log \kappa(n)}{n} = \frac{\log \frac{1}{\beta_{\mathsf{A}_{\mathcal{K}}}(n)}}{[k+1]n}$$

$$\leq \frac{\log \frac{n^{\xi/k}}{\kappa'}}{[k+1]n}$$
$$= \frac{\xi \log n}{k[k+1]n} - \frac{\log \kappa'}{[k+1]n} \tag{79}$$

for all sufficiently large n. This implies that $\boldsymbol{\kappa}$ satisfies (78). In the following, κ denotes $\kappa(n)$.

From (78), we have the fact that there is a γ such that $\gamma > 0$ and

$$\eta_{\mathcal{X}_j|\mathcal{U}}(\gamma|\gamma) + \frac{\log \kappa}{n} \leq \varepsilon_j \tag{80}$$

$$[k+3]\gamma + \sum_{j\in\mathcal{K}} \iota_{\mathcal{X}_j|\mathcal{U}}(\gamma|\gamma) \leq \sum_{j\in\mathcal{K}} \varepsilon_j \tag{81}$$

for all $j \in \mathcal{K}$ and sufficiently large n.

When $\boldsymbol{u} \in \mathcal{T}_{U,\gamma}$, we have

$$\begin{aligned}
|\mathcal{T}_{X_j|U,\gamma}(\boldsymbol{u})| &\geq 2^{n[H(X_j|U) - \eta_{\mathcal{X}_j|\mathcal{U}}(\gamma|\gamma)]} \\
&\geq \kappa 2^{n[H(X_j|U)-\varepsilon_j]} \\
&= \kappa 2^{n[r_j+R_j]} \\
&= \kappa |\mathrm{Im}\mathcal{A}_j||\mathrm{Im}\mathcal{A}'_j| \\
&\geq \kappa |\mathrm{Im}\widehat{\mathcal{A}}_j|
\end{aligned} \tag{82}$$

for all $j \in \mathcal{K}$ and sufficiently large n, where the first inequality comes from Lemma 13, the second inequality comes from (80), the first equality comes from (31) and (32), and the last inequality comes from the fact that $\mathrm{Im}\widehat{\mathcal{A}}_j \subset \mathrm{Im}\mathcal{A}_j \times \mathrm{Im}\mathcal{A}'_j$.

This implies that for all $j \in \mathcal{K}$ and sufficiently large n there is $\mathcal{T}_j(\boldsymbol{u}) \subset \mathcal{T}_{X_j|U,\gamma}(\boldsymbol{u})$ such that

$$\kappa \leq \frac{|\mathcal{T}_j(\boldsymbol{u})|}{|\mathrm{Im}\widehat{\mathcal{A}}_j|} \leq 2\kappa. \tag{83}$$

for all $\boldsymbol{u}$. We assume that $\mathcal{T}_j(\boldsymbol{u})$ is constructed by selecting $|\mathcal{T}_j(\boldsymbol{u})|$ elements in the ascending order regarding the value $D(\nu_{\boldsymbol{x}_j|\boldsymbol{u}}\|\mu_{X_j|U}|\nu_{\boldsymbol{u}})$.

Let $\boldsymbol{m}_{\mathcal{K}} \in \mathcal{M}_{\mathcal{K}}$ be private messages. Let $\boldsymbol{x}_{\mathcal{K}}$ be channel inputs, where $\boldsymbol{x}_j \in \mathcal{X}_j^n$ is defined as

$$\boldsymbol{x}_j \equiv \widehat{g}_{A_j A'_j}(\boldsymbol{a}_j, \boldsymbol{m}_j|\boldsymbol{u}) \quad \text{for each } j \in \mathcal{K}.$$

Let $\boldsymbol{y} \in \mathcal{Y}^n$ be a channel output. We define

$$\begin{aligned}
\mathcal{S}_j &\equiv \{(\boldsymbol{m}_{\mathcal{K}}, \boldsymbol{y}) : \boldsymbol{x}_j \in \mathcal{T}_j(\boldsymbol{u}) \text{ and } \boldsymbol{y} \in \mathcal{Y}^n\} \\
\mathcal{S}_{k+1} &\equiv \left\{ (\boldsymbol{m}_{\mathcal{K}}, \boldsymbol{y}) : I(\boldsymbol{x}_{\mathcal{K}}|\boldsymbol{u}) < \gamma + \sum_{j\in\mathcal{K}} \left[\iota_{\mathcal{X}_j|\mathcal{U}}(\gamma|\gamma) + \varepsilon_j\right] \text{ and } \boldsymbol{y} \in \mathcal{Y}^n \right\} \\
\mathcal{S}_{k+2} &\equiv \left\{ (\boldsymbol{m}_{\mathcal{K}}, \boldsymbol{y}) : \boldsymbol{y} \in \mathcal{T}_{Y|UX_{\mathcal{K}},\gamma}(\boldsymbol{u}, \boldsymbol{x}_{\mathcal{K}}) \right\} \\
\mathcal{S}_{k+3} &\equiv \left\{ (\boldsymbol{m}_{\mathcal{K}}, \boldsymbol{y}) : \widehat{g}_{A_{\mathcal{K}}}(\boldsymbol{a}_{\mathcal{K}}|\boldsymbol{y}, \boldsymbol{u}) = \boldsymbol{x}_{\mathcal{K}} \right\},
\end{aligned}$$

where $j \in \mathcal{K}$ and

$$I(\boldsymbol{x}_{\mathcal{K}}|\boldsymbol{u}) \equiv \sum_{j\in\mathcal{K}} H(\boldsymbol{x}_j|\boldsymbol{u}) - H(\boldsymbol{x}_{\mathcal{K}}|\boldsymbol{u}).$$

We assign equation numbers to the conditions

$$\boldsymbol{u} \in \mathcal{T}_{U,\gamma} \tag{84}$$

$$\boldsymbol{x}_j \in \mathcal{T}_j(\boldsymbol{u}) \subset \mathcal{T}_{X_j|U,\gamma}(\boldsymbol{u}) \quad \text{for all } j \in \mathcal{K} \tag{85}$$

$$I(\boldsymbol{x}_{\mathcal{K}}|\boldsymbol{u}) < \gamma + \sum_{j\in\mathcal{K}} \left[\iota_{\mathcal{X}_j|\mathcal{U}}(\gamma|\gamma) + \varepsilon_j\right] \tag{86}$$

$$\boldsymbol{y} \in \mathcal{T}_{Y|UX_{\mathcal{K}},\gamma}(\boldsymbol{u}, \boldsymbol{x}_{\mathcal{K}}) \tag{87}$$

$$\widehat{g}_{A_{\mathcal{K}}}(\boldsymbol{a}_{\mathcal{K}}|\boldsymbol{y}, \boldsymbol{u}) \neq \boldsymbol{x}_{\mathcal{K}}, \tag{88}$$

which are referred later. Since the j-th message $\boldsymbol{m}_j$ satisfies $A'_j\boldsymbol{x}_j = \boldsymbol{m}_j$, the decoder can recover message $\boldsymbol{m}_{\mathcal{K}}$ when decoding the channel input $\boldsymbol{x}_{\mathcal{K}}$ is successful. This implies that the decoding error probability is upper bounded by

$$\begin{aligned}
&\mathrm{Error}(A_{\mathcal{K}}, A'_{\mathcal{K}}, \boldsymbol{a}_{\mathcal{K}}, \boldsymbol{u}) \\
&\le \sum_{j\in\mathcal{K}} p_{M_{\mathcal{K}}Y}(\mathcal{S}_j^c) + p_{M_{\mathcal{K}}Y}\left(\left[\cap_{j=1}^{k}\mathcal{S}_j\right] \cap \mathcal{S}_{k+1}^c\right) + p_{M_{\mathcal{K}}Y}(\mathcal{S}_{k+2}^c) + p_{M_{\mathcal{K}}Y}\left(\left[\cap_{j=1}^{k+2}\mathcal{S}_j\right] \cap \mathcal{S}_{k+3}^c\right).
\end{aligned} \tag{89}$$

Remark 4: The condition (86) was unnecessary in the proof of the conventional random coding argument because $\boldsymbol{x}_{\mathcal{K}} \in \mathcal{T}_{X_{\mathcal{K}|U},\gamma}(\boldsymbol{u})$ was naturally satisfied by generating codewords independently at random for a given $\boldsymbol{u} \in \mathcal{U}^n$. On the other hand, (86) is necessary in our proof because (84) and (85) may not imply $(\boldsymbol{u}, \boldsymbol{x}_{\mathcal{K}}) \in \mathcal{T}_{UX_{\mathcal{K}},\gamma'}$ for an appropriate $\gamma' > 0$. This is the difference from the conventional proof. It should be noted that (84)–(87) implies $(\boldsymbol{u}, \boldsymbol{x}_{\mathcal{K}}, \boldsymbol{y}) \in \mathcal{T}_{UX_{\mathcal{K}}Y,\gamma'}$, where $\gamma' > 0$ will be specified later.

In the following we evaluate the average error probability

$$\begin{aligned}
&E_{\widehat{\mathsf{A}}_{\mathcal{K}}\mathbf{a}_{\mathcal{K}}U^n}\left[\mathrm{Error}(\mathsf{A}_{\mathcal{K}}, \mathsf{A}'_{\mathcal{K}}, \mathbf{a}_{\mathcal{K}}, U^n)\right] \\
&\le E_{\widehat{\mathsf{A}}_{\mathcal{K}}\mathbf{a}_{\mathcal{K}}}\left[\sum_{\boldsymbol{u}\in\mathcal{T}_{U,\gamma}} \mu_U(\boldsymbol{u})\mathrm{Error}(\mathsf{A}_{\mathcal{K}}, \mathsf{A}'_{\mathcal{K}}, \mathbf{a}_{\mathcal{K}}, \boldsymbol{u})\right] + \mu_U\left([\mathcal{T}_{U,\gamma}]^c\right) \\
&\le \sum_{j\in\mathcal{K}} E_{\widehat{\mathsf{A}}_{\mathcal{K}}\mathbf{a}_{\mathcal{K}}}\left[\sum_{\boldsymbol{u}\in\mathcal{T}_{U,\gamma}} \mu_U(\boldsymbol{u})p_{M_{\mathcal{K}}Y}(\mathcal{S}_j^c)\right] + E_{\widehat{\mathsf{A}}_{\mathcal{K}}\mathbf{a}_{\mathcal{K}}}\left[\sum_{\boldsymbol{u}\in\mathcal{T}_{U,\gamma}} \mu_U(\boldsymbol{u})p_{M_{\mathcal{K}}Y}\left(\left[\cap_{j=1}^{k}\mathcal{S}_j\right] \cap \mathcal{S}_{k+1}^c\right)\right] \\
&\quad + E_{\widehat{\mathsf{A}}_{\mathcal{K}}\mathbf{a}_{\mathcal{K}}}\left[\sum_{\boldsymbol{u}\in\mathcal{T}_{U,\gamma}} \mu_U(\boldsymbol{u})p_{M_{\mathcal{K}}Y}(\mathcal{S}_{k+2}^c)\right] + E_{\widehat{\mathsf{A}}_{\mathcal{K}}\mathbf{a}_{\mathcal{K}}}\left[\sum_{\boldsymbol{u}\in\mathcal{T}_{U,\gamma}} \mu_U(\boldsymbol{u})p_{M_{\mathcal{K}}Y}\left(\left[\cap_{j=1}^{k+2}\mathcal{S}_j\right] \cap \mathcal{S}_{k+3}^c\right)\right] \\
&\quad + \mu_U\left([\mathcal{T}_{U,\gamma}]^c\right),
\end{aligned} \tag{90}$$

over random variables $\widehat{\mathsf{A}}_{\mathcal{K}}$, $\mathbf{a}_{\mathcal{K}}$, and U^n. The last term on the right hand side of (90) is evaluated as

$$\begin{aligned}
\mu_U\left([\mathcal{T}_{U,\gamma}]^c\right) &\le 2^{-n[\gamma-\lambda_{\mathcal{U}}]} \\
&\le \frac{\delta}{k+4}
\end{aligned} \tag{91}$$

for all $\delta > 0$ and all sufficiently large n, where the first inequality comes from Lemma 12. In the following, let

$$\widehat{\boldsymbol{a}}_j \equiv (\boldsymbol{a}_j, \boldsymbol{m}_j) \quad \text{for each } j \in \mathcal{K}.$$

We assume that the distribution of $\widehat{\boldsymbol{a}}_j$ is uniform on $\mathrm{Im}\widehat{\mathcal{A}}_j$ for all $j \in \mathcal{K}$, and random variables $\{\widehat{\mathsf{A}}_j, \mathbf{a}_j, M_j\}_{j\in\mathcal{K}}$ and U^n are mutually independent. In the following, we use the fact that $\widehat{g}_{A_jA'_j}(\boldsymbol{a}_j, \boldsymbol{m}_j|\boldsymbol{u}) \notin \mathcal{T}_j(\boldsymbol{u})$ implies

$\mathcal{T}_j(\boldsymbol{u}) \cap \mathcal{C}_{\widehat{A}_j}(\widehat{\boldsymbol{a}}_j) = \emptyset$, which is shown by contradiction as follows. Let us assume that $\boldsymbol{x}_j \equiv \widehat{g}_{A_j A'_j}(\boldsymbol{a}_j, \boldsymbol{m}_j|\boldsymbol{u}) \notin \mathcal{T}_j(\boldsymbol{u})$ and $\mathcal{T}_j(\boldsymbol{u}) \cap \mathcal{C}_{\widehat{A}_j}(\widehat{\boldsymbol{a}}_j) \neq \emptyset$. Then there is $\boldsymbol{x}'_j \in \mathcal{T}_j(\boldsymbol{u}) \cap \mathcal{C}_{\widehat{A}_j}(\widehat{\boldsymbol{a}}_j)$. From the definition of $\widehat{g}_{A_j A'_j}$, we have

$$D(\nu_{\boldsymbol{x}_j|\boldsymbol{u}} \| \mu_{X_j|U} | \nu_{\boldsymbol{u}}) \leq D(\nu_{\boldsymbol{x}'_j|\boldsymbol{u}} \| \mu_{X_j|U} | \nu_{\boldsymbol{u}}). \tag{92}$$

On the other hand, from the construction of $\mathcal{T}_j(\boldsymbol{u})$, we have the fact that $\boldsymbol{x}''_j \in \mathcal{T}_j(\boldsymbol{u})$ if $\boldsymbol{x}'_j \in \mathcal{T}_j(\boldsymbol{u})$ and

$$D(\nu_{\boldsymbol{x}''_j|\boldsymbol{u}} \| \mu_{X_j|U} | \nu_{\boldsymbol{u}}) \leq D(\nu_{\boldsymbol{x}'_j|\boldsymbol{u}} \| \mu_{X_j|U} | \nu_{\boldsymbol{u}}).$$

From this fact and (92), we have the fact that $\boldsymbol{x}_j \in \mathcal{T}_j(\boldsymbol{u})$, which contradicts the assumption $\boldsymbol{x}_j \equiv \widehat{g}_{A_j A'_j}(\boldsymbol{a}_j, \boldsymbol{m}_j|\boldsymbol{u}) \notin \mathcal{T}_j(\boldsymbol{u})$. First, we evaluate $E_{\widehat{\mathsf{A}}_{\mathcal{K}} \mathbf{a}_{\mathcal{K}}} \left[\sum_{\boldsymbol{u} \in \mathcal{T}_{U,\gamma}} \mu_U(\boldsymbol{u}) p_{M_{\mathcal{K}} Y}(\mathcal{S}_j^c) \right]$. From Lemma 2 and (83), we have

$$\begin{aligned}
E_{\widehat{\mathsf{A}}_{\mathcal{K}} \mathbf{a}_{\mathcal{K}}} \left[\sum_{\boldsymbol{u} \in \mathcal{T}_{U,\gamma}} \mu_U(\boldsymbol{u}) p_{M_{\mathcal{K}} Y}(\mathcal{S}_j^c) \right] &= \sum_{\boldsymbol{u} \in \mathcal{T}_{U,\gamma}} \mu_U(\boldsymbol{u}) p_{\widehat{\mathsf{A}}_j \mathbf{a}_j M_j} \left(\left\{ (A_j, A'_j, \boldsymbol{a}_j, \boldsymbol{m}_j) : \widehat{g}_{A_j A'_j}(\boldsymbol{a}_j, \boldsymbol{m}_j|\boldsymbol{u}) \notin \mathcal{T}_j(\boldsymbol{u}) \right\} \right) \\
&\leq \sum_{\boldsymbol{u} \in \mathcal{T}_{U,\gamma}} \mu_U(\boldsymbol{u}) p_{\widehat{\mathsf{A}}_j \widehat{\mathbf{a}}_j} \left(\left\{ (\widehat{A}_j, \widehat{\boldsymbol{a}}_j) : \mathcal{T}_j(\boldsymbol{u}) \cap \mathcal{C}_{\widehat{A}_j}(\widehat{\boldsymbol{a}}_j) = \emptyset \right\} \right) \\
&\leq \sum_{\boldsymbol{u} \in \mathcal{T}_{U,\gamma}} \mu_U(\boldsymbol{u}) \left[\alpha_{\widehat{\mathsf{A}}_j} - 1 + \frac{|\mathrm{Im} \widehat{\mathcal{A}}_j| \left[\beta_{\widehat{\mathsf{A}}_j} + 1 \right]}{|\mathcal{T}_j(\boldsymbol{u})|} \right] \\
&\leq \sum_{\boldsymbol{u} \in \mathcal{T}_{U,\gamma}} \mu_U(\boldsymbol{u}) \left[\alpha_{\widehat{\mathsf{A}}_j} - 1 + \frac{\beta_{\widehat{\mathsf{A}}_j} + 1}{\kappa} \right] \\
&\leq \frac{\delta}{k+4}
\end{aligned} \tag{93}$$

for all $\delta > 0$ and sufficiently large n, where the first inequality comes from the fact that $\widehat{g}_{A_j A'_j}(\boldsymbol{a}_j, \boldsymbol{m}_j|\boldsymbol{u}) \notin \mathcal{T}_j(\boldsymbol{u})$ implies $\mathcal{T}_j(\boldsymbol{u}) \cap \mathcal{C}_{\widehat{A}_j}(\widehat{\boldsymbol{a}}_j) = \emptyset$, and the last inequality comes from (76) and the fact that $\alpha_{\widehat{\mathsf{A}}_j} \to 1$ and $\beta_{\widehat{\mathsf{A}}_j} \to 0$ as $n \to \infty$.

Next, we evaluate the second term on the right hand side of (90). Assume that $\boldsymbol{x}_{\mathcal{K}}$ satisfies (84), (85), and

$$I(\boldsymbol{x}_{\mathcal{K}}|\boldsymbol{u}) \geq \gamma + \sum_{j \in \mathcal{K}} \left[\iota_{\mathcal{X}_j|\mathcal{U}}(\gamma|\gamma) + \varepsilon_j \right].$$

Then, from Lemma 11, we have

$$|H(\boldsymbol{x}_j|\boldsymbol{u}) - H(X_j|U)| < \iota_{\mathcal{X}_j|\mathcal{U}}(\gamma|\gamma) \quad \text{for all } j \in \mathcal{K}.$$

We have

$$\begin{aligned}
H(\boldsymbol{x}_{\mathcal{K}}|\boldsymbol{u}) &= \sum_{j \in \mathcal{K}} H(\boldsymbol{x}_j|\boldsymbol{u}) - I(\boldsymbol{x}_{\mathcal{K}}|\boldsymbol{u}) \\
&\leq \sum_{j \in \mathcal{K}} \left[H(X_j|U) + \iota_{\mathcal{X}_j|\mathcal{U}}(\gamma|\gamma) \right] - \left[\gamma + \sum_{j \in \mathcal{K}} \left[\iota_{\mathcal{X}_j|\mathcal{U}}(\gamma|\gamma) + \varepsilon_j \right] \right] \\
&= \sum_{j \in \mathcal{K}} [r_j + R_j] - \gamma,
\end{aligned} \tag{94}$$

where the last equality comes from (30). Since $\boldsymbol{x}_j \in \mathcal{C}_{A_j A'_j}(\boldsymbol{a}_j, \boldsymbol{m}_j)$ for all $j \in \mathcal{K}$, we have

$$\boldsymbol{x}_{\mathcal{K}} \in \mathcal{G}(\boldsymbol{u}) \cap \mathcal{C}_{\widehat{A}_{\mathcal{K}}}(\widehat{\boldsymbol{a}}_{\mathcal{K}}),$$

where $\mathcal{G}(\boldsymbol{u}) \subset \times_{j\in\mathcal{K}} \mathcal{X}_j^n$ is defined as

$$\mathcal{G}(\boldsymbol{u}) \equiv \left\{ \boldsymbol{x}_{\mathcal{K}} : H(\boldsymbol{x}_{\mathcal{K}}|\boldsymbol{u}) < \sum_{j\in\mathcal{K}} [r_j + R_j] - \gamma \right\}.$$

This implies that

$$\mathcal{G}(\boldsymbol{u}) \cap \mathcal{C}_{\widehat{A}_{\mathcal{K}}}(\widehat{\boldsymbol{a}}_{\mathcal{K}}) \neq \emptyset.$$

Then we have

$$\begin{aligned}
p_{\widehat{\mathsf{A}}_{\mathcal{K}}\widehat{\mathbf{a}}_{\mathcal{K}}}\left(\left\{(\widehat{A}_{\mathcal{K}}, \widehat{\boldsymbol{a}}_{\mathcal{K}}) : \mathcal{G}(\boldsymbol{u}) \cap \mathcal{C}_{\widehat{A}_{\mathcal{K}}}(\widehat{\boldsymbol{a}}_{\widetilde{\mathcal{K}}}) \neq \emptyset\right\}\right) &\leq \sum_{\boldsymbol{x}_{\mathcal{K}}\in\mathcal{G}(\boldsymbol{u})} p_{\widehat{\mathsf{A}}_{\mathcal{K}}\widehat{\mathbf{a}}_{\mathcal{K}}}\left(\left\{(\widehat{A}_{\mathcal{K}}, \widehat{\boldsymbol{a}}_{\mathcal{K}}) : \widehat{A}_j \boldsymbol{x}_j = \widehat{\boldsymbol{a}}_j \text{for all } j \in \mathcal{K}\right\}\right) \\
&= \sum_{\boldsymbol{x}_{\mathcal{K}}\in\mathcal{G}(\boldsymbol{u})} \sum_{\widehat{A}_{\mathcal{K}}, \widehat{\boldsymbol{a}}_{\mathcal{K}}} \prod_{j\in\mathcal{K}} p_{\widehat{\mathsf{A}}_j\widehat{\mathbf{a}}_j}(\widehat{A}_j, \widehat{\boldsymbol{a}}_j)\chi\left(\widehat{A}_j \boldsymbol{x}_j = \widehat{\boldsymbol{a}}_j\right) \\
&= \sum_{\boldsymbol{x}_{\mathcal{K}}\in\mathcal{G}(\boldsymbol{u})} \prod_{j\in\mathcal{K}} \left[\sum_{\widehat{A}_j, \widehat{\boldsymbol{a}}_j} p_{\widehat{\mathsf{A}}_j\widehat{\mathbf{a}}_j}(\widehat{A}_j, \widehat{\boldsymbol{a}}_j)\chi\left(\widehat{A}_j \boldsymbol{x}_j = \widehat{\boldsymbol{a}}_j\right)\right] \\
&= \sum_{\boldsymbol{x}_{\mathcal{K}}\in\mathcal{G}(\boldsymbol{u})} \frac{1}{\prod_{j\in\mathcal{K}} |\mathrm{Im}\widehat{\mathcal{A}}_j|} \\
&= \frac{|\mathcal{G}(\boldsymbol{u})|}{\prod_{j\in\mathcal{K}} |\mathrm{Im}\widehat{\mathcal{A}}_j|} \\
&\leq \frac{2^{n\left[\sum_{j\in\mathcal{K}}[r_j+R_j]-\gamma+\lambda_{\mathcal{X}_{\mathcal{K}}}\right]}}{\prod_{j\in\mathcal{K}} |\mathrm{Im}\widehat{\mathcal{A}}_j|} \\
&= 2^{-n[\gamma-\lambda_{\mathcal{X}_{\mathcal{K}}}]}, \qquad (95)
\end{aligned}$$

where the second inequality comes from Lemma 8. This implies that

$$\begin{aligned}
E_{\widehat{\mathsf{A}}_{\mathcal{K}}\mathbf{a}_{\mathcal{K}}}\left[\sum_{\boldsymbol{u}\in\mathcal{T}_{U,\gamma}} \mu_U(\boldsymbol{u}) p_{M_{\mathcal{K}}Y}([\cap_{j=1}^k \mathcal{S}_j] \cap \mathcal{S}_{k+1}^c)\right] &\leq \sum_{\boldsymbol{u}\in\mathcal{T}_{U,\gamma}} \mu_U(\boldsymbol{u}) p_{\widehat{\mathsf{A}}_{\mathcal{K}}\widehat{\mathbf{a}}_{\mathcal{K}}}\left(\left\{(\widehat{A}_{\mathcal{K}}, \widehat{\boldsymbol{a}}_{\mathcal{K}}) : \mathcal{G}(\boldsymbol{u}) \cap \mathcal{C}_{\widehat{A}_{\mathcal{K}}}(\widehat{\boldsymbol{a}}_{\widetilde{\mathcal{K}}}) \neq \emptyset\right\}\right) \\
&\leq \sum_{\boldsymbol{u}\in\mathcal{T}_{U,\gamma}} \mu_U(\boldsymbol{u}) 2^{-n[\gamma-\lambda_{\mathcal{X}_{\mathcal{K}}}]} \\
&\leq \frac{\delta}{k+4} \qquad (96)
\end{aligned}$$

for all $\delta > 0$ and sufficiently large n, where the last inequality comes from the fact that $\lambda_{\mathcal{X}_{\mathcal{K}}} \to 0$ as $n \to \infty$.

Next, we evaluate the third term on the right hand side of (90). Let $\boldsymbol{X}_{\mathcal{K}} \equiv \{\widehat{g}_{\mathsf{A}_j\mathsf{A}'_j}(\mathbf{a}_j, M_j|\boldsymbol{u})\}_{j\in\mathcal{K}}$. Then we have

$$\begin{aligned}
\mu_{Y|X_{\mathcal{K}}}\left(\left[\mathcal{T}_{Y|UX_{\mathcal{K}},\gamma}(\boldsymbol{u}, \boldsymbol{X}_{\mathcal{K}})\right]^c | \boldsymbol{X}_{\mathcal{K}}\right) &= \mu_{Y|U^n X_{\mathcal{K}}}\left(\left[\mathcal{T}_{Y|UX_{\mathcal{K}},\gamma}(\boldsymbol{u}, \boldsymbol{X}_{\mathcal{K}})\right]^c | \boldsymbol{u}, \boldsymbol{X}_{\mathcal{K}}\right) \\
&\leq 2^{-n[\gamma-\lambda_{\mathcal{U}\mathcal{X}_{\mathcal{K}}\mathcal{Y}}]}. \qquad (97)
\end{aligned}$$

from Lemma 12. This implies that

$$\begin{aligned}
E_{\widehat{\mathsf{A}}_{\mathcal{K}}\mathbf{a}_{\mathcal{K}}}\left[\sum_{\boldsymbol{u}\in\mathcal{T}_{U,\gamma}} \mu_U(\boldsymbol{u}) p_{M_{\mathcal{K}}Y}(\mathcal{S}_{k+2}^c)\right] &= E_{\widehat{\mathsf{A}}_{\mathcal{K}}\widehat{\mathbf{a}}_{\mathcal{K}}}\left[\sum_{\boldsymbol{u}\in\mathcal{T}_{U,\gamma}} \mu_U(\boldsymbol{u}) \mu_{Y|X_{\mathcal{K}}}\left(\left[\mathcal{T}_{Y|UX_{\mathcal{K}},\gamma}(\boldsymbol{u}, \boldsymbol{X}_{\mathcal{K}})\right]^c | \boldsymbol{X}_{\mathcal{K}}\right)\right] \\
&\leq 2^{-n[\gamma-\lambda_{\mathcal{U}\mathcal{X}_{\mathcal{K}}\mathcal{Y}}]} \\
&\leq \frac{\delta}{k+4} \qquad (98)
\end{aligned}$$

for all $\delta > 0$ and sufficiently large n, where the last inequality comes from the fact that $\lambda_{\mathcal{U}\mathcal{X}_\mathcal{K}\mathcal{Y}} \to 0$ as $n \to \infty$.

Next, we evaluate the fourth term on the right hand side of (90). In the following, we assume that (84)–(87) and

$$\widehat{g}_{A_\mathcal{K}}(\boldsymbol{a}_\mathcal{K}|\boldsymbol{y},\boldsymbol{u}) \neq \boldsymbol{x}_\mathcal{K}.x$$

Then, from (11), we have

$$\begin{aligned}
D(\nu_{\boldsymbol{u}\boldsymbol{x}_\mathcal{K}\boldsymbol{y}} \parallel \mu_{UX_\mathcal{K}Y}) &= \sum_{u,x_\mathcal{K},y} \nu_{\boldsymbol{u}\boldsymbol{x}_\mathcal{K}\boldsymbol{y}}(u,x_\mathcal{K},y) \log \frac{\nu_{\boldsymbol{u}\boldsymbol{x}_\mathcal{K}\boldsymbol{y}}(u,x_\mathcal{K},y)}{\mu_{UX_\mathcal{K}Y}(u,x_\mathcal{K},y)} \\
&= \sum_{u,x_\mathcal{K},y} \nu_{\boldsymbol{u}\boldsymbol{x}_\mathcal{K}\boldsymbol{y}}(u,x_\mathcal{K},y) \log \frac{\nu_{\boldsymbol{y}|\boldsymbol{u}\boldsymbol{x}_\mathcal{K}}(y|u,x_\mathcal{K})}{\mu_{Y|UX_\mathcal{K}}(y|u,x_\mathcal{K})} + \sum_{j\in\mathcal{K}} \sum_{u,x_j} \nu_{\boldsymbol{u}\boldsymbol{x}_j}(u,x_j) \log \frac{\nu_{\boldsymbol{x}_j|\boldsymbol{u}}(x_j|u)}{\mu_{X_j|U}(x_j|u)} \\
&\quad + \sum_u \nu_{\boldsymbol{u}}(u) \log \frac{\nu_{\boldsymbol{u}}(u)}{\mu_U(u)} + \sum_{x_\mathcal{K}} \nu_{\boldsymbol{u}\boldsymbol{x}_\mathcal{K}}(u,x_\mathcal{K}) \log \frac{\nu_{\boldsymbol{x}_\mathcal{K}|\boldsymbol{u}}(x_\mathcal{K}|u)}{\prod_{j\in\mathcal{K}} \nu_{\boldsymbol{x}_j|\boldsymbol{u}}(x_j|u)} \\
&= D(\nu_{\boldsymbol{y}|\boldsymbol{u}\boldsymbol{x}_\mathcal{K}} \parallel \mu_{Y|UX_\mathcal{K}}|\nu_{\boldsymbol{u}\boldsymbol{x}_\mathcal{K}}) + \sum_{j\in\mathcal{K}} D(\nu_{\boldsymbol{x}_j|\boldsymbol{u}} \parallel \mu_{X_j|U}|\nu_{\boldsymbol{u}}) + D(\nu_{\boldsymbol{u}}\|\mu_U) + I(\boldsymbol{x}_\mathcal{K}|\boldsymbol{u}) \\
&< [k+2]\gamma + \gamma + \sum_{j\in\mathcal{K}} \left[\iota_{\mathcal{X}_j|\mathcal{U}}(\gamma|\gamma) + \varepsilon_j\right] \\
&\leq 2\sum_{j\in\mathcal{K}} \varepsilon_j \qquad (99)
\end{aligned}$$

where the last inequality comes from (81). This implies that

$$(\boldsymbol{u},\boldsymbol{x}_\mathcal{K},\boldsymbol{y}) \in \mathcal{T}_{X_\mathcal{K}Y,\gamma'},$$

where γ' is defined as

$$\gamma' \equiv 2\sum_{j\in\mathcal{K}} \varepsilon_j.$$

Since $\widehat{g}_{A_\mathcal{K}}(\boldsymbol{a}_\mathcal{K}|\boldsymbol{y},\boldsymbol{u}) \neq \boldsymbol{x}_\mathcal{K}$, there is $\boldsymbol{x}'_\mathcal{K} \in \mathcal{C}_{A_\mathcal{K}}(\boldsymbol{a}_\mathcal{K})$ such that $\boldsymbol{x}'_\mathcal{K} \neq \boldsymbol{x}_\mathcal{K}$ and $(\boldsymbol{u},\boldsymbol{x}'_\mathcal{K},\boldsymbol{y}) \in \mathcal{T}_{UX_\mathcal{K}Y,\gamma'}$. This implies that

$$[\mathcal{G}(\boldsymbol{u},\boldsymbol{y}) \setminus \{\boldsymbol{x}_\mathcal{K}\}] \cap \mathcal{C}_{A_\mathcal{K}}(A_\mathcal{K}\boldsymbol{x}_\mathcal{K}) \neq \emptyset,$$

where

$$\mathcal{G}(\boldsymbol{u},\boldsymbol{y}) \equiv \{\boldsymbol{x}_\mathcal{K} : (\boldsymbol{u},\boldsymbol{x}_\mathcal{K},\boldsymbol{y}) \in \mathcal{T}_{UX_\mathcal{K}Y,\gamma'}\}.$$

From Lemma 9, we have the fact that

$$\mathcal{G}(\boldsymbol{u},\boldsymbol{y}) \subset \mathcal{T}_{X_\mathcal{K}|UY,\gamma'}(\boldsymbol{u},\boldsymbol{y})$$

and $\boldsymbol{x}_\mathcal{K} \in \mathcal{G}(\boldsymbol{u},\boldsymbol{y})$ implies $(\boldsymbol{u},\boldsymbol{y}) \in \mathcal{T}_{UY,\gamma'}$. Then, from Lemma 13, we have

$$\begin{aligned}
|\mathcal{G}_{\mathcal{K}|\mathcal{K}^c}(\boldsymbol{u},\boldsymbol{y})| &\equiv |\mathcal{G}(\boldsymbol{u},\boldsymbol{y})| \\
&\leq |\mathcal{T}_{X_\mathcal{K}|UY,\gamma'}(\boldsymbol{u},\boldsymbol{y})| \\
&\leq 2^{n[H(X_\mathcal{K}|UY)+\eta_{\mathcal{X}_\mathcal{K}|\mathcal{U}\mathcal{Y}}(\gamma'|\gamma')]}. \qquad (100)
\end{aligned}$$

For each non-empty set $\mathcal{J} \subsetneq \mathcal{K}$, let

$$\mathcal{G}_{\mathcal{X}_{\mathcal{J}^c}}(\boldsymbol{u},\boldsymbol{y}) \equiv \left\{\boldsymbol{x}_{\mathcal{J}^c} : \boldsymbol{x}_\mathcal{K} \in \mathcal{G}(\boldsymbol{u},\boldsymbol{y}) \text{ for some } \boldsymbol{x}_\mathcal{J} \in \mathcal{X}^n_\mathcal{J}\right\}$$

$$\mathcal{G}_{\mathcal{X}_\mathcal{J}|\mathcal{X}_{\mathcal{J}^c}}(\boldsymbol{u},\boldsymbol{x}_{\mathcal{J}^c},\boldsymbol{y}) \equiv \{\boldsymbol{x}_\mathcal{J} : \boldsymbol{x}_\mathcal{K} \in \mathcal{G}(\boldsymbol{u},\boldsymbol{y})\}.$$

Then, from Lemma 9, we have the fact that $\boldsymbol{x}_{\mathcal{J}^c} \in \mathcal{G}_{\mathcal{X}_{\mathcal{J}^c}}(\boldsymbol{u},\boldsymbol{y})$ implies $(\boldsymbol{u},\boldsymbol{x}_{\mathcal{J}^c},\boldsymbol{y}) \in \mathcal{T}_{UX_{\mathcal{J}^c}Y,\gamma'}$ and

$$\mathcal{G}_{\mathcal{X}_{\mathcal{J}}|\mathcal{X}_{\mathcal{J}^c}}(\boldsymbol{u},\boldsymbol{x}_{\mathcal{J}^c},\boldsymbol{y}) \subset \mathcal{T}_{X_{\mathcal{J}}|UX_{\mathcal{J}^c}Y,\gamma'}(\boldsymbol{u},\boldsymbol{x}_{\mathcal{J}^c},\boldsymbol{y})$$

for every non-empty set $\mathcal{J} \subsetneq \mathcal{K}$. We have

$$\begin{aligned}
|\mathcal{G}_{\mathcal{J}|\mathcal{J}^c}(\boldsymbol{u},\boldsymbol{y})| &\equiv \max_{\boldsymbol{x}_{\mathcal{J}^c}\in\mathcal{G}_{\mathcal{X}_{\mathcal{J}^c}}(\boldsymbol{u},\boldsymbol{y})} \left|\mathcal{G}_{\mathcal{X}_{\mathcal{J}}|\mathcal{X}_{\mathcal{J}^c}}(\boldsymbol{u},\boldsymbol{x}_{\mathcal{J}^c},\boldsymbol{y})\right| \\
&\le \max_{(\boldsymbol{u},\boldsymbol{x}_{\mathcal{J}^c},\boldsymbol{y})\in\mathcal{T}_{UX_{\mathcal{J}^c}Y,\gamma'}} \left|\mathcal{T}_{X_{\mathcal{J}}|UX_{\mathcal{J}^c}Y,\gamma'}(\boldsymbol{u},\boldsymbol{x}_{\mathcal{J}^c},\boldsymbol{y})\right| \\
&\le 2^{n[H(X_{\mathcal{J}}|U,X_{\mathcal{J}^c},Y)+\eta_{\mathcal{X}_{\mathcal{J}}|\mathcal{U}\mathcal{X}_{\mathcal{J}^c}\mathcal{Y}}(\gamma'|\gamma')]} \\
&\le 2^{n\left[H(X_{\mathcal{J}}|U,X_{\mathcal{J}^c},Y)+\eta_{\mathcal{X}_{\mathcal{K}}|\mathcal{U}\mathcal{Y}}(\gamma'|\gamma')\right]} \qquad (101)
\end{aligned}$$

for every non-empty set $\mathcal{J} \subsetneq \mathcal{K}$, where the second inequality comes from Lemma 13. Then, from (100), (101), and Lemma 4, we have

$$\begin{aligned}
E_{\mathsf{A}_{\mathcal{K}}}\left[\chi(\widehat{g}_{\mathsf{A}_{\mathcal{K}}}(\mathsf{A}_{\mathcal{K}}\boldsymbol{x}_{\mathcal{K}}|\boldsymbol{y},\boldsymbol{u}) \neq \boldsymbol{x}_{\mathcal{K}})\right] &\le p_{\mathsf{A}_{\mathcal{K}}}\left(\{A_{\mathcal{K}} : [\mathcal{G}(\boldsymbol{u},\boldsymbol{y}) \setminus \{\boldsymbol{x}_{\mathcal{K}}\}] \cap \mathcal{C}_{A_{\mathcal{K}}}(A_{\mathcal{K}}\boldsymbol{x}_{\mathcal{K}}) \neq \emptyset\}\right) \\
&\le \sum_{\substack{\mathcal{J}\subset\mathcal{K}\\ \mathcal{J}\neq\emptyset}} \frac{2^{n\left[H(X_{\mathcal{J}}|UX_{\mathcal{J}^c},Y)+\eta_{\mathcal{X}_{\mathcal{K}}|\mathcal{U}\mathcal{Y}}(\gamma'|\gamma')\right]}\alpha_{\mathsf{A}_{\mathcal{J}}}\left[\beta_{\mathsf{A}_{\mathcal{J}^c}}+1\right]}{\prod_{j\in\mathcal{J}}|\mathrm{Im}\mathcal{A}_j|} + \beta_{\mathsf{A}_{\mathcal{K}}} \qquad (102)
\end{aligned}$$

for all $(\boldsymbol{u},\boldsymbol{x}_{\mathcal{K}},\boldsymbol{y}) \in \mathcal{T}_{UX_{\mathcal{K}}Y,\gamma'}$. Then we have

$$\begin{aligned}
&E_{\mathsf{A}_{\mathcal{K}}\mathsf{A}'_{\mathcal{K}}\mathbf{a}_{\mathcal{K}}}\left[\sum_{\boldsymbol{u}\in\mathcal{T}_{U,\gamma}} \mu_U(\boldsymbol{u}) p_{M_{\mathcal{K}}Y}([\cap_{j=1}^{k+2}\mathcal{S}_j]\cap\mathcal{S}_{k+3}^c)\right] \\
&\le E_{\mathsf{A}_{\mathcal{K}}\mathbf{a}_{\mathcal{K}}M_{\mathcal{K}}}\left[\sum_{\boldsymbol{u}\in\mathcal{T}_{U,\gamma}} \mu_U(\boldsymbol{u}) \sum_{\boldsymbol{x}_{\mathcal{K}}\in\mathcal{T}_{\mathcal{K}}(\boldsymbol{u})}\left[\prod_{j\in\mathcal{K}}\chi(\widehat{g}_{\mathsf{A}_j\mathsf{A}'_j}(\mathbf{a}_j,M_j|\boldsymbol{u})=\boldsymbol{x}_j)\right]\right. \\
&\qquad\qquad \left.\sum_{\boldsymbol{y}\in\mathcal{T}_{Y|X_{\mathcal{K}},\gamma}(\boldsymbol{x}_{\mathcal{K}})} \mu_{Y|X_{\mathcal{K}}}(\boldsymbol{y}|\boldsymbol{x}_{\mathcal{K}})\chi(\widehat{g}_{\mathsf{A}_{\mathcal{K}}}(\mathbf{a}_{\mathcal{K}}|\boldsymbol{y},\boldsymbol{u})\neq\boldsymbol{x}_{\mathcal{K}})\right] \\
&\le \sum_{\substack{\boldsymbol{u}\in\mathcal{T}_{U,\gamma}\\ \boldsymbol{x}_{\mathcal{K}}\in\mathcal{T}_{\mathcal{K}}(\boldsymbol{u})\\ \boldsymbol{y}\in\mathcal{T}_{Y|X_{\mathcal{K}},\gamma}(\boldsymbol{x}_{\mathcal{K}})}} \mu_U(\boldsymbol{u})\mu_{Y|X_{\mathcal{K}}}(\boldsymbol{y}|\boldsymbol{x}_{\mathcal{K}}) \\
&\qquad\qquad \cdot E_{\mathsf{A}_{\mathcal{K}}}\left[\chi(\widehat{g}_{\mathsf{A}_{\mathcal{K}}}(\mathsf{A}_{\mathcal{K}}\boldsymbol{x}_{\mathcal{K}}|\boldsymbol{y},\boldsymbol{u})\neq\boldsymbol{x}_{\mathcal{K}})\prod_{j\in\mathcal{K}}E_{\mathbf{a}_j}\left[\chi(\mathsf{A}_j\boldsymbol{x}_j=\mathbf{a}_j)\right]E_{\mathsf{A}'_jM_j}\left[\chi(\mathsf{A}'_j\boldsymbol{x}_j=M_j)\right]\right] \\
&= \frac{1}{\prod_{j\in\mathcal{K}}|\mathrm{Im}\widehat{A}_j|} \sum_{\substack{\boldsymbol{u}\in\mathcal{T}_{U,\gamma}\\ \boldsymbol{x}_{\mathcal{K}}\in\mathcal{T}_{\mathcal{K}}(\boldsymbol{u})\\ \boldsymbol{y}\in\mathcal{T}_{Y|X_{\mathcal{K}},\gamma}(\boldsymbol{x}_{\mathcal{K}})}} \mu_U(\boldsymbol{u})\mu_{Y|X_{\mathcal{K}}}(\boldsymbol{y}|\boldsymbol{x}_{\mathcal{K}})E_{\mathsf{A}_{\mathcal{K}}}\left[\chi(\widehat{g}_{\mathsf{A}_{\mathcal{K}}}(\mathsf{A}_{\mathcal{K}}\boldsymbol{x}_{\mathcal{K}}|\boldsymbol{y},\boldsymbol{u})\neq\boldsymbol{x}_{\mathcal{K}})\right] \\
&\le \frac{1}{\prod_{j\in\mathcal{K}}|\mathrm{Im}\widehat{A}_j|} \sum_{\substack{\boldsymbol{u}\in\mathcal{T}_{U,\gamma}\\ \boldsymbol{x}_{\mathcal{K}}\in\mathcal{T}_{\mathcal{K}}(\boldsymbol{u})\\ \boldsymbol{y}\in\mathcal{T}_{Y|X_{\mathcal{K}},\gamma}(\boldsymbol{x}_{\mathcal{K}})}} \mu_U(\boldsymbol{u})\mu_{Y|X_{\mathcal{K}}}(\boldsymbol{y}|\boldsymbol{x}_{\mathcal{K}}) \\
&\qquad\qquad \cdot \left[\sum_{\substack{\mathcal{J}\subset\mathcal{K}\\ \mathcal{J}\neq\emptyset}} \frac{2^{n\left[H(X_{\mathcal{J}}|U,X_{\mathcal{J}^c},Y)+\eta_{\mathcal{X}_{\mathcal{K}}|\mathcal{U}\mathcal{Y}}(\gamma'|\gamma')\right]}\alpha_{\mathsf{A}_{\mathcal{J}}}[\beta_{\mathsf{A}_{\mathcal{J}^c}}+1]}{\prod_{j\in\mathcal{J}}|\mathrm{Im}\mathcal{A}_j|} + \beta_{\mathsf{A}_{\mathcal{K}}}\right]
\end{aligned}$$

$$\leq 2^k \kappa^k \left[\sum_{\substack{\mathcal{J} \subset \mathcal{K} \\ \mathcal{J} \neq \emptyset}} 2^{-n\left[\sum_{j \in \mathcal{J}} r_j - H(X_{\mathcal{J}}|U, X_{\mathcal{J}^c}, Y) - \eta_{\mathcal{X}_{\mathcal{K}}|\mathcal{U}\mathcal{Y}}(\gamma'|\gamma')\right]} \alpha_{\mathsf{A}_{\mathcal{J}}} [\beta_{\mathsf{A}_{\mathcal{J}^c}} + 1] + \beta_{\mathsf{A}_{\mathcal{K}}} \right]$$

$$\leq \frac{\delta}{k+4} \tag{103}$$

for all $\delta > 0$ and all sufficiently large n, where $\mathcal{T}_{\mathcal{K}}(\boldsymbol{u})$ is defined as

$$\mathcal{T}_{\mathcal{K}}(\boldsymbol{u}) \equiv \prod_{j \in \mathcal{K}} \mathcal{T}_j(\boldsymbol{u}),$$

the equality comes from Lemma 5 that appears in Appendix A, the third inequality comes from (102), the fourth inequality comes from (31) and (83), and the last inequality comes from (26), (27), (35), and (77).

Finally, from (90)–(93), (96), (98), and (103), we have the fact that for all $\delta > 0$ and all sufficiently large n there are $\{A_j, A'_j, \boldsymbol{a}_j\}_{j \in \mathcal{K}}$, and $\boldsymbol{u}$ satisfying $A_j \in \mathcal{A}$, $A'_j \in \mathcal{A}'_j$, $\boldsymbol{a}_j \in \mathrm{Im}\mathcal{A}_j$, $\boldsymbol{u} \in \mathcal{U}^n$ and (36). ■

B. Proof of Theorem 3

We can prove the theorem similarly to the proof of Theorem 1.

In the following, we assume that ensembles $(\boldsymbol{\mathcal{A}}_j, \boldsymbol{p}_{\mathsf{A}_j})$ and $(\boldsymbol{\mathcal{A}}'_j, \boldsymbol{p}_{\mathsf{A}'_j})$ have a hash property for all $j \in \widetilde{\mathcal{K}}$. Similarly to the proof of Theorem 1, we define an ensemble $(\widehat{\boldsymbol{\mathcal{A}}}_j, \boldsymbol{p}_{\widehat{\mathsf{A}}_j})$ and $(\boldsymbol{\alpha}_{\widehat{\mathsf{A}}_j}, \boldsymbol{\beta}_{\widehat{\mathsf{A}}_j})$ for each $j \in \widetilde{\mathcal{K}}$. Then we have the fact that $(\widehat{\boldsymbol{\mathcal{A}}}_j, \boldsymbol{p}_{\widehat{\mathsf{A}}_j})$ has a $(\boldsymbol{\alpha}_{\widehat{\mathsf{A}}_j}, \boldsymbol{\beta}_{\widehat{\mathsf{A}}_j})$-hash property and there is a sequence $\boldsymbol{\kappa} \equiv \{\kappa(n)\}_{n=1}^{\infty}$ such that

$$\lim_{n \to \infty} \kappa(n) = \infty \tag{104}$$

$$\lim_{n \to \infty} [\kappa(n)]^3 \beta_{\mathsf{A}_{\widetilde{\mathcal{K}}}}(n) = 0 \tag{105}$$

$$\lim_{n \to \infty} \frac{\log \kappa(n)}{n} = 0. \tag{106}$$

From (106), we have the fact that there is a γ such that $\gamma > 0$ and

$$\eta_{\mathcal{X}_0}(\gamma) + \frac{\log \kappa}{n} \leq \varepsilon_0 \tag{107}$$

$$\eta_{\mathcal{X}_j|\mathcal{X}_0}(\gamma|\gamma) + \frac{\log \kappa}{n} \leq \varepsilon_j \tag{108}$$

for all $j \in \widetilde{\mathcal{K}}$ and all sufficiently large n and

$$5\gamma + \sum_{j \in \widetilde{\mathcal{K}}} \iota_j(\gamma) \leq \sum_{j \in \widetilde{\mathcal{K}}} \varepsilon_j, \tag{109}$$

where $\iota_j(\gamma)$ is defined by

$$\iota_j(\gamma) \equiv \begin{cases} \iota_{\mathcal{X}_0}(\gamma) & \text{if } j = 0 \\ \iota_{\mathcal{X}_j|\mathcal{X}_0}(\gamma|\gamma) & \text{if } j \in \mathcal{K}. \end{cases}$$

Similarly to the proof of (83), from (107), we have the fact that there is a set $\mathcal{T}_0$ such that $\mathcal{T}_0 \subset \mathcal{T}_{X_0,\gamma}$ and

$$\kappa \leq \frac{|\mathcal{T}_0|}{|\mathrm{Im}\widehat{\mathcal{A}}_0|} \leq 2\kappa. \tag{110}$$

We assume that $\mathcal{T}_0$ is constructed by selecting $|\mathcal{T}_0|$ elements in the ascending order regarding the value $D(\nu_{\boldsymbol{x}_0} \| \mu_{X_0})$. Furthermore, from (108), we have the fact that for all $\boldsymbol{x}_0 \in \mathcal{X}_0^n$ and all $j \in \mathcal{K}$ there is a

set $\mathcal{T}_j(\boldsymbol{x}_0)$ such that $\mathcal{T}_j(\boldsymbol{x}_0) \subset \mathcal{T}_{X_j|X_0,\gamma}(\boldsymbol{x}_0)$ and

$$\kappa \le \frac{|\mathcal{T}_j(\boldsymbol{x}_0)|}{|\mathrm{Im}\widehat{\mathcal{A}}_1|} \le 2\kappa. \tag{111}$$

We assume that $\mathcal{T}_j(\boldsymbol{x}_0)$ is constructed by selecting $|\mathcal{T}_j(\boldsymbol{x}_0)|$ elements in the ascending order regarding the value $D(\nu_{\boldsymbol{x}_j|\boldsymbol{x}_0}\|\mu_{X_j|X_0}|\nu_{\boldsymbol{x}_0})$.

Now we prove the theorem. Let $\boldsymbol{m}_0 \in \mathcal{M}_0$ be a common message and $\boldsymbol{m}_1 \in \mathcal{M}_1$ and $\boldsymbol{m}_2 \in \mathcal{M}_2$ be private messages. Let $\boldsymbol{x}_0 \in \mathcal{X}_0^n$ be defined as

$$\boldsymbol{x}_0 \equiv \widehat{g}_{A_0A_0'}(\boldsymbol{a}_0, \boldsymbol{m}_0).$$

Let $(\boldsymbol{x}_1, \boldsymbol{x}_2)$ be channel inputs, where $\boldsymbol{x}_j \in \mathcal{X}_j^n$ is defined by

$$\boldsymbol{x}_j \equiv \widehat{g}_{A_jA_j'}(\boldsymbol{a}_j, \boldsymbol{m}_j|\boldsymbol{x}_0) \quad \text{for each } j \in \mathcal{K}.$$

Let $\boldsymbol{y} \in \mathcal{Y}^n$ be the channel output. We define

$$\begin{aligned}
\mathcal{S}_0 &\equiv \left\{(\boldsymbol{m}_{\widetilde{\mathcal{K}}}, \boldsymbol{y}) : \boldsymbol{x}_0 \in \mathcal{T}_0 \subset \mathcal{T}_{X_0,\gamma} \text{ and } \boldsymbol{y} \in \mathcal{Y}^n\right\} \\
\mathcal{S}_j &\equiv \left\{(\boldsymbol{m}_{\widetilde{\mathcal{K}}}, \boldsymbol{y}) : \boldsymbol{x}_j \in \mathcal{T}_j(\boldsymbol{x}_0) \subset \mathcal{T}_{X_j|X_0,\gamma}(\boldsymbol{x}_0) \text{ and } \boldsymbol{y} \in \mathcal{Y}^n\right\} \\
\mathcal{S}_3 &\equiv \left\{(\boldsymbol{m}_{\widetilde{\mathcal{K}}}, \boldsymbol{y}) : I(\boldsymbol{x}_1;\boldsymbol{x}_2|\boldsymbol{x}_0) < \gamma + \sum_{j\in\widetilde{\mathcal{K}}} [\iota_j(\gamma) + \varepsilon_j] \text{ and } \boldsymbol{y} \in \mathcal{Y}^n\right\} \\
\mathcal{S}_4 &\equiv \left\{(\boldsymbol{m}_{\widetilde{\mathcal{K}}}, \boldsymbol{y}) : \boldsymbol{y} \in \mathcal{T}_{Y|X_{\widetilde{\mathcal{K}}},\gamma}(\boldsymbol{x}_{\widetilde{\mathcal{K}}})\right\} \\
\mathcal{S}_5 &\equiv \left\{(\boldsymbol{m}_{\widetilde{\mathcal{K}}}, \boldsymbol{y}) : \widehat{g}_{A_{\widetilde{\mathcal{K}}}}(\boldsymbol{a}_{\widetilde{\mathcal{K}}}|\boldsymbol{y}) = \boldsymbol{x}_{\widetilde{\mathcal{K}}}\right\},
\end{aligned}$$

where $j \in \mathcal{K}$ and

$$\begin{aligned}
I(\boldsymbol{x}_1;\boldsymbol{x}_2|\boldsymbol{x}_0) &\equiv \sum_{j\in\{1,2\}} H(\boldsymbol{x}_j|\boldsymbol{x}_0) - H(\boldsymbol{x}_{\{1,2\}}|\boldsymbol{x}_0) \\
&= H(\boldsymbol{x}_1|\boldsymbol{x}_0) + H(\boldsymbol{x}_2|\boldsymbol{x}_0) - H(\boldsymbol{x}_1,\boldsymbol{x}_2|\boldsymbol{x}_0).
\end{aligned} \tag{112}$$

The error probability is upper bounded by

$$\begin{aligned}
&\mathrm{Error}(A_{\widetilde{\mathcal{K}}}, A_{\widetilde{\mathcal{K}}}', \boldsymbol{a}_{\widetilde{\mathcal{K}}}) \\
&\le p_{M_{\widetilde{\mathcal{K}}}Y}(\mathcal{S}_0^c) + \sum_{j\in\mathcal{K}} p_{M_{\widetilde{\mathcal{K}}}Y}(\mathcal{S}_0 \cap \mathcal{S}_j^c) + p_{M_{\widetilde{\mathcal{K}}}Y}\left(\left[\cap_{j=0}^2 \mathcal{S}_j\right] \cap \mathcal{S}_3^c\right) + p_{M_{\widetilde{\mathcal{K}}}Y}(\mathcal{S}_4^c) + p_{M_{\widetilde{\mathcal{K}}}Y}\left(\left[\cap_{j=0}^4 \mathcal{S}_j\right] \cap \mathcal{S}_5^c\right),
\end{aligned} \tag{113}$$

We assign equation numbers to the conditions

$$\boldsymbol{x}_0 \in \mathcal{T}_0 \subset \mathcal{T}_{X_0,\gamma} \tag{114}$$

$$\boldsymbol{x}_j \in \mathcal{T}_j(\boldsymbol{x}_0) \subset \mathcal{T}_{X_j|X_0,\gamma}(\boldsymbol{x}_0) \quad \text{for all } j \in \mathcal{K} \equiv 1, 2 \tag{115}$$

$$I(\boldsymbol{x}_1;\boldsymbol{x}_2|\boldsymbol{x}_0) < \gamma + \sum_{j\in\widetilde{\mathcal{K}}} [\iota_j(\gamma) + \varepsilon_j] \tag{116}$$

$$\boldsymbol{y} \in \mathcal{T}_{Y|X_{\widetilde{\mathcal{K}}},\gamma}(\boldsymbol{x}_{\widetilde{\mathcal{K}}}) \tag{117}$$

$$\widehat{g}_{A_{\widetilde{\mathcal{K}}}}(\boldsymbol{a}_{\widetilde{\mathcal{K}}}|\boldsymbol{y}) = \boldsymbol{x}_{\widetilde{\mathcal{K}}} \tag{118}$$

which are referred later. In comparison with the conventional superposition coding, the condition (114) corresponds to an event where the function $\widehat{g}_{A_0A_0'}$ finds a 'good' cloud center $\boldsymbol{x}_0$ and the condition (115) corresponds to an event where the function $\widehat{g}_{A_jA_j'}$ finds a 'good' satellite $\boldsymbol{x}_j$ for all $j \in \mathcal{K}$, where 'good' means that they are (conditionally) typical sequences. When (114)–(117) are satisfied, we have the fact that $(\boldsymbol{x}_0, \boldsymbol{x}_1, \boldsymbol{x}_2)$ is jointly typical. It should be noted that (116) was unnecessary in the proof of the conventional superposition coding because the joint typicality of $(\boldsymbol{x}_0, \boldsymbol{x}_1, \boldsymbol{x}_2)$ was naturally satisfied by generating codewords at random.

In the following, let

$$\widehat{\boldsymbol{a}}_j \equiv (\boldsymbol{a}_j, \boldsymbol{m}_j) \quad \text{for each } j \in \widetilde{\mathcal{K}}.$$

We assume that the distribution of $\widehat{\boldsymbol{a}}_j$ is uniform on $\mathrm{Im}\widehat{\mathcal{A}}_j$ for all $j \in \widetilde{\mathcal{K}}$, and $\{\widehat{\mathsf{A}}_j, \mathbf{a}_j, M_j\}_{j\in\widetilde{\mathcal{K}}}$ are mutually independent.

First, we evaluate $E_{\widehat{\mathsf{A}}_{\widetilde{\mathcal{K}}}\mathbf{a}_{\widetilde{\mathcal{K}}}}\left[p_{M_{\widetilde{\mathcal{K}}}Y}(\mathcal{S}_0^c)\right]$. From Lemma 2 and (110), we have

$$\begin{aligned}
E_{\widehat{\mathsf{A}}_{\widetilde{\mathcal{K}}}\mathbf{a}_{\widetilde{\mathcal{K}}}}\left[p_{M_{\widetilde{\mathcal{K}}}Y}(\mathcal{S}_0^c)\right] &= p_{\widehat{\mathsf{A}}_0\mathbf{a}_0M_0}\left(\left\{(A_0, A_0', \boldsymbol{a}_0, \boldsymbol{m}_0) : \widehat{g}_{A_0A_0'}(\boldsymbol{a}_0, \boldsymbol{m}_0) \notin \mathcal{T}_0\right\}\right) \\
&\le p_{\widehat{\mathsf{A}}_0\widehat{\mathbf{a}}_0}\left(\left\{(\widehat{A}_0, \widehat{\boldsymbol{a}}_0) : \mathcal{T}_0 \cap \mathcal{C}_{\widehat{A}_0}(\widehat{\boldsymbol{a}}_0) = \emptyset\right\}\right) \\
&\le \alpha_{\widehat{\mathsf{A}}_0} - 1 + \frac{|\mathrm{Im}\widehat{\mathcal{A}}_0|\left[\beta_{\widehat{\mathsf{A}}_0} + 1\right]}{|\mathcal{T}_0|} \\
&\le \alpha_{\widehat{\mathsf{A}}_0} - 1 + \frac{\beta_{\widehat{\mathsf{A}}_0} + 1}{\kappa} \\
&\le \frac{\delta}{6}
\end{aligned} \tag{119}$$

for all $\delta > 0$ and sufficiently large n, where the last inequality comes from (104) and the fact that $\alpha_{\widehat{\mathsf{A}}_0} \to 1$ and $\beta_{\widehat{\mathsf{A}}_0} \to 0$ as $n \to \infty$.

Next, we evaluate $E_{\widehat{\mathsf{A}}_{\widetilde{\mathcal{K}}}\mathbf{a}_{\widetilde{\mathcal{K}}}}\left[p_{M_{\widetilde{\mathcal{K}}}Y}(\mathcal{S}_0 \cap \mathcal{S}_j^c)\right]$. From Lemma 2 and (111), we have

$$\begin{aligned}
&E_{\widehat{\mathsf{A}}_{\widetilde{\mathcal{K}}}\mathbf{a}_{\widetilde{\mathcal{K}}}}\left[p_{M_{\widetilde{\mathcal{K}}}Y}(\mathcal{S}_0 \cap \mathcal{S}_j^c)\right] \\
&= \sum_{\widehat{A}_0, \widehat{\mathbf{a}}_0} p_{\widehat{\mathsf{A}}_0\widehat{\mathbf{a}}_0}(\widehat{A}_0, \widehat{\boldsymbol{a}}_0) \sum_{\boldsymbol{x}_0 \in \mathcal{T}_0} \chi\left(\widehat{g}_{\widehat{A}_0}(\widehat{\boldsymbol{a}}_0) = \boldsymbol{x}_0\right) p_{\widehat{\mathsf{A}}_j\widehat{\mathbf{a}}_j}\left(\left\{(\widehat{A}_j, \widehat{\boldsymbol{a}}_j) : \widehat{g}_{\widehat{A}_j}(\widehat{\boldsymbol{a}}_j|\boldsymbol{x}_0) \notin \mathcal{T}_j(\boldsymbol{x}_0)\right\}\right) \\
&\le \sum_{\widehat{A}_0, \widehat{\mathbf{a}}_0} p_{\widehat{\mathsf{A}}_0\widehat{\mathbf{a}}_0}(\widehat{A}_0, \widehat{\boldsymbol{a}}_0) \sum_{\boldsymbol{x}_0 \in \mathcal{T}_0} \chi\left(\widehat{g}_{\widehat{A}_0}(\widehat{\boldsymbol{a}}_0) = \boldsymbol{x}_0\right) p_{\widehat{\mathsf{A}}_j\widehat{\mathbf{a}}_j}\left(\left\{(\widehat{A}_j, \widehat{\boldsymbol{a}}_j) : \mathcal{T}_j(\boldsymbol{x}_0) \cap \mathcal{C}_{\widehat{A}_j}(\widehat{\boldsymbol{a}}_j) = \emptyset\right\}\right) \\
&\le \sum_{\widehat{A}_0, \widehat{\mathbf{a}}_0} p_{\widehat{\mathsf{A}}_0\widehat{\mathbf{a}}_0}(\widehat{A}_0, \widehat{\boldsymbol{a}}_0) \sum_{\boldsymbol{x}_0 \in \mathcal{T}_0} \chi\left(\widehat{g}_{\widehat{A}_0}(\widehat{\boldsymbol{a}}_0) = \boldsymbol{x}_0\right) \left[\alpha_{\widehat{\mathsf{A}}_j} - 1 + \frac{|\mathrm{Im}\widehat{\mathcal{A}}_j|\left[\beta_{\widehat{\mathsf{A}}_j} + 1\right]}{|\mathcal{T}_j(\boldsymbol{x}_0)|}\right] \\
&\le \alpha_{\widehat{\mathsf{A}}_j} - 1 + \frac{\beta_{\widehat{\mathsf{A}}_j} + 1}{\kappa} \\
&\le \frac{\delta}{6}
\end{aligned} \tag{120}$$

for all $\delta > 0$ and sufficiently large n, where the last inequality comes from (104) and the fact that $\alpha_{\widehat{\mathsf{A}}_j} \to 1$ and $\beta_{\widehat{\mathsf{A}}_j} \to 0$ as $n \to \infty$.

Next, we evaluate $E_{\widehat{\mathsf{A}}_{\widetilde{\mathcal{K}}}\mathbf{a}_{\widetilde{\mathcal{K}}}}\left[p_{M_{\widetilde{\mathcal{K}}}Y}\left(\left[\cap_{j=0}^2 \mathcal{S}_j\right] \cap \mathcal{S}_3^c\right)\right]$. Assume that $(\boldsymbol{x}_0, \boldsymbol{x}_1, \boldsymbol{x}_2)$ satisfies (114), (115) and

$$I(\boldsymbol{x}_1; \boldsymbol{x}_2|\boldsymbol{x}_0) \ge \gamma + \sum_{j\in\widetilde{\mathcal{K}}} \left[\iota_j(\gamma) + \varepsilon_j\right].$$

Then, from Lemma 11, we have

$$|H(\boldsymbol{x}_0) - H(X_0)| < \iota_{\mathcal{X}_0}(\gamma)$$
$$|H(\boldsymbol{x}_1|\boldsymbol{x}_0) - H(X_1|X_0)| < \iota_{\mathcal{X}_1|\mathcal{X}_0}(\gamma|\gamma)$$
$$|H(\boldsymbol{x}_2|\boldsymbol{x}_0) - H(X_2|X_0)| < \iota_{\mathcal{X}_2|\mathcal{X}_0}(\gamma|\gamma).$$

Then we have

$$\begin{aligned} H(\boldsymbol{x}_0, \boldsymbol{x}_1, \boldsymbol{x}_2) &= H(\boldsymbol{x}_0) + H(\boldsymbol{x}_1|\boldsymbol{x}_0) + H(\boldsymbol{x}_2|\boldsymbol{x}_0) - I(\boldsymbol{x}_1; \boldsymbol{x}_2|\boldsymbol{x}_0) \\ &\le H(X_0) + H(X_1|X_0) + H(X_2|X_0) - \left[\gamma + \sum_{j\in\widetilde{\mathcal{K}}} [\iota_j(\gamma) + \varepsilon_j]\right] \\ &= \sum_{j\in\widetilde{\mathcal{K}}} [r_j + R_j] - \gamma, \end{aligned} \tag{121}$$

where the last equality comes from (51)–(53). Since $\boldsymbol{x}_j \in \mathcal{C}_{A_jB_j}(\boldsymbol{a}_j, \boldsymbol{m}_j)$ for all $j \in \widetilde{\mathcal{K}}$, we have

$$(\boldsymbol{x}_0, \boldsymbol{x}_1, \boldsymbol{x}_2) \in \mathcal{G} \cap \mathcal{C}_{\widehat{A}_{\widetilde{\mathcal{K}}}}(\widehat{\boldsymbol{a}}_{\widetilde{\mathcal{K}}}),$$

where $\mathcal{G} \subset \mathcal{X}_0^n \times \mathcal{X}_1^n \times \mathcal{X}_2^n$ is defined as

$$\mathcal{G} \equiv \left\{ (\boldsymbol{x}_0, \boldsymbol{x}_1, \boldsymbol{x}_2) : H(\boldsymbol{x}_0, \boldsymbol{x}_1, \boldsymbol{x}_2) < \sum_{j\in\widetilde{\mathcal{K}}} [r_j + R_j] - \gamma \right\}.$$

This implies that

$$\mathcal{G} \cap \mathcal{C}_{\widehat{A}_{\widetilde{\mathcal{K}}}}(\widehat{\boldsymbol{a}}_{\widetilde{\mathcal{K}}}) \neq \emptyset.$$

Similarly to the proof of (96), we have

$$\begin{aligned} E_{\widehat{\mathsf{A}}_{\widetilde{\mathcal{K}}}\mathbf{a}_{\widetilde{\mathcal{K}}}}\left[p_{M_{\widetilde{\mathcal{K}}}Y}\left([\cap_{j=0}^{2}\mathcal{S}_j] \cap \mathcal{S}_3^c\right)\right] &\le p_{\widehat{\mathsf{A}}_{\widetilde{\mathcal{K}}}\mathbf{a}_{\widetilde{\mathcal{K}}}}\left(\left\{(\widehat{A}_{\widetilde{\mathcal{K}}}, \widehat{\boldsymbol{a}}_{\widetilde{\mathcal{K}}}) : \mathcal{G} \cap \mathcal{C}_{\widehat{A}_{\widetilde{\mathcal{K}}}}(\widehat{\boldsymbol{a}}_{\widetilde{\mathcal{K}}}) \neq \emptyset\right\}\right) \\ &\le \sum_{\boldsymbol{x}_{\widetilde{\mathcal{K}}}\in\mathcal{G}} p_{\widehat{\mathsf{A}}_{\widetilde{\mathcal{K}}}\mathbf{a}_{\widetilde{\mathcal{K}}}}\left(\left\{(\widehat{A}_{\widetilde{\mathcal{K}}}, \widehat{\boldsymbol{a}}_{\widetilde{\mathcal{K}}}) : \widehat{A}_j\boldsymbol{x}_j = \widehat{\boldsymbol{a}}_j \text{ for all } j \in \widetilde{\mathcal{K}}\right\}\right) \\ &= \frac{|\mathcal{G}|}{\prod_{j\in\widetilde{\mathcal{K}}} |\mathrm{Im}\widehat{\mathcal{A}}_j|} \\ &\le \frac{2^{n\left[\sum_{j\in\widetilde{\mathcal{K}}}[r_j+R_j]-\gamma+\lambda_{\mathcal{X}_{\widetilde{\mathcal{K}}}}\right]}}{\prod_{j\in\widetilde{\mathcal{K}}} |\mathrm{Im}\widehat{\mathcal{A}}_j|} \\ &= 2^{-n[\gamma-\lambda_{\mathcal{X}_{\widetilde{\mathcal{K}}}}]} \\ &\le \frac{\delta}{6} \end{aligned} \tag{122}$$

for all $\delta > 0$ and all sufficiently large n.

Next, we evaluate $E_{\widehat{\mathsf{A}}_{\widetilde{\mathcal{K}}}\mathbf{a}_{\widetilde{\mathcal{K}}}}\left[p_{M_{\widetilde{\mathcal{K}}}Y}(\mathcal{S}_4^c)\right]$. Similarly to the proof of (98), we have

$$\begin{aligned} E_{\widehat{\mathsf{A}}_{\widetilde{\mathcal{K}}}\mathbf{a}_{\widetilde{\mathcal{K}}}}\left[p_{M_{\widetilde{\mathcal{K}}}Y}(\mathcal{S}_4^c)\right] &= E_{\widehat{\mathsf{A}}_{\widetilde{\mathcal{K}}}\widehat{\mathbf{a}}_{\widetilde{\mathcal{K}}}}\left[\mu_{Y|X_1X_2}\left(\left[\mathcal{T}_{Y|X_{\widetilde{\mathcal{K}}},\gamma}(\boldsymbol{X}_{\widetilde{\mathcal{K}}})\right]^c \middle| \boldsymbol{X}_1, \boldsymbol{X}_2\right)\right] \\ &= E_{\widehat{\mathsf{A}}_{\widetilde{\mathcal{K}}}\widehat{\mathbf{a}}_{\widetilde{\mathcal{K}}}}\left[\mu_{Y|X_{\widetilde{\mathcal{K}}}}\left(\left[\mathcal{T}_{Y|X_{\widetilde{\mathcal{K}}},\gamma}(\boldsymbol{X}_{\widetilde{\mathcal{K}}})\right]^c \middle| \boldsymbol{X}_{\widetilde{\mathcal{K}}}\right)\right] \\ &\le 2^{-n[\gamma-\lambda_{\mathcal{X}_{\widetilde{\mathcal{K}}}\mathcal{Y}}]} \\ &\le \frac{\delta}{6} \end{aligned} \tag{123}$$

for all $\delta > 0$ and all sufficiently large n, where $\boldsymbol{X}_{\widetilde{\mathcal{K}}} \equiv \{\boldsymbol{X}_i\}_{i\in\widetilde{\mathcal{K}}}$ is defined by

$$\boldsymbol{X}_0 \equiv \widehat{g}_{A_0A_0'}(\mathbf{a}_0, M_0)$$

$$\boldsymbol{X}_j \equiv \widehat{g}_{A_jA_j'}(\mathbf{a}_j, M_j|\boldsymbol{X}_0) \quad \text{for each } j \in \mathcal{K}.$$

Next, we evaluate $E_{\widehat{\mathsf{A}}_{\widetilde{\mathcal{K}}}\mathbf{a}_{\widetilde{\mathcal{K}}}}\left[p_{M_{\widetilde{\mathcal{K}}}Y}\left(\left[\cap_{j=0}^4\mathcal{S}_j\right]\cap\mathcal{S}_5^c\right)\right]$. In the following, we assume (114)–(117) and

$$g_{A_0A_1A_2}(\boldsymbol{a}_0, \boldsymbol{a}_1, \boldsymbol{a}_2|\boldsymbol{y}) \neq (\boldsymbol{x}_0, \boldsymbol{x}_1, \boldsymbol{x}_2).$$

Similarly to the proof of (99), we have

$$\begin{aligned} D(\nu_{\boldsymbol{x}_{\widetilde{\mathcal{K}}}\boldsymbol{y}} \parallel \mu_{X_{\widetilde{\mathcal{K}}}Y}) &= D(\nu_{\boldsymbol{y}|\boldsymbol{x}_{\widetilde{\mathcal{K}}}} \parallel \mu_{Y|X_{\widetilde{\mathcal{K}}}}|\nu_{\boldsymbol{x}_{\widetilde{\mathcal{K}}}}) + \sum_{j\in\{0,1\}} D(\nu_{\boldsymbol{x}_j|\boldsymbol{x}_0} \parallel \mu_{X_j|X_0}|\nu_{\boldsymbol{x}_0}) + D(\nu_{\boldsymbol{x}_0} \parallel \mu_{X_0}) + I(\boldsymbol{x}_1;\boldsymbol{x}_2|\boldsymbol{x}_0) \\ &< 5\gamma + \sum_{j\in\widetilde{\mathcal{K}}}\left[\iota_j(\gamma) + \varepsilon_j\right] \\ &\leq 2\sum_{j\in\widetilde{\mathcal{K}}}\varepsilon_j, \end{aligned} \tag{124}$$

where the last inequality comes from (109). This implies that

$$(\boldsymbol{x}_0, \boldsymbol{x}_1, \boldsymbol{x}_2, \boldsymbol{y}) \in \mathcal{T}_{X_{\widetilde{\mathcal{K}}}Y,\gamma'}.$$

where γ' is defined as

$$\gamma' \equiv 2\sum_{j\in\widetilde{\mathcal{K}}}\varepsilon_j.$$

Since $\widehat{g}_{A_0A_1A_2}(\boldsymbol{a}_0, \boldsymbol{a}_1, \boldsymbol{a}_2|\boldsymbol{y}) \neq (\boldsymbol{x}_0, \boldsymbol{x}_1, \boldsymbol{x}_2)$, there is $(\boldsymbol{x}_0', \boldsymbol{x}_1', \boldsymbol{x}_2') \in \mathcal{C}_{A_{\widetilde{\mathcal{K}}}}(\boldsymbol{a}_{\widetilde{\mathcal{K}}})$ such that $(\boldsymbol{x}_0', \boldsymbol{x}_1', \boldsymbol{x}_2') \neq (\boldsymbol{x}_0, \boldsymbol{x}_1, \boldsymbol{x}_2)$ and $(\boldsymbol{x}_0', \boldsymbol{x}_1', \boldsymbol{x}_2', \boldsymbol{y}) \in \mathcal{T}_{X_{\widetilde{\mathcal{K}}}Y,\gamma'}$. This implies that

$$\left[\mathcal{G}(\boldsymbol{y}) \setminus \{\boldsymbol{x}_{\widetilde{\mathcal{K}}}\}\right] \cap \mathcal{C}_{A_{\widetilde{\mathcal{K}}}}(A_{\widetilde{\mathcal{K}}}\boldsymbol{x}_{\widetilde{\mathcal{K}}}) \neq \emptyset,$$

where

$$\mathcal{G}(\boldsymbol{y}) \equiv \left\{(\boldsymbol{x}_0, \boldsymbol{x}_1, \boldsymbol{x}_2) : (\boldsymbol{x}_0, \boldsymbol{x}_1, \boldsymbol{x}_2, \boldsymbol{y}) \in \mathcal{T}_{X_{\widetilde{\mathcal{K}}}Y,\gamma'}\right\}.$$

From Lemma 9, we have the fact that

$$\mathcal{G}(\boldsymbol{y}) \subset \mathcal{T}_{X_{\widetilde{\mathcal{K}}}|Y,\gamma'}(\boldsymbol{y})$$

and $(\boldsymbol{x}_0, \boldsymbol{x}_1, \boldsymbol{x}_2) \in \mathcal{G}(\boldsymbol{y})$ implies $\boldsymbol{y} \in \mathcal{T}_{Y,\gamma'}$. Then, from Lemma 13, we have

$$\begin{aligned} |\mathcal{G}_{\widetilde{\mathcal{K}}|\widetilde{\mathcal{K}}^c}(\boldsymbol{y})| &\equiv |\mathcal{G}(\boldsymbol{y})| \\ &\leq |\mathcal{T}_{X_{\widetilde{\mathcal{K}}}|Y,\gamma'}(\boldsymbol{y})| \\ &\leq 2^{n[H(X_{\widetilde{\mathcal{K}}}|Y)+\eta_{\mathcal{X}_{\widetilde{\mathcal{K}}}|\mathcal{Y}}(\gamma'|\gamma')]}. \end{aligned} \tag{125}$$

For each non-empty set $\mathcal{J} \subsetneq \widetilde{\mathcal{K}}$, let

$$\mathcal{G}_{\mathcal{X}_{\mathcal{J}^c}}(\boldsymbol{y}) \equiv \left\{\boldsymbol{x}_{\mathcal{J}^c} : \boldsymbol{x}_{\widetilde{\mathcal{K}}} \in \mathcal{G}(\boldsymbol{y}) \text{ for some } \boldsymbol{x}_{\mathcal{J}} \in \mathcal{X}_{\mathcal{J}}^n\right\}$$

$$\mathcal{G}_{\mathcal{X}_{\mathcal{J}}|\mathcal{X}_{\mathcal{J}^c}}(\boldsymbol{x}_{\mathcal{J}^c}, \boldsymbol{y}) \equiv \left\{\boldsymbol{x}_{\mathcal{J}} : \boldsymbol{x}_{\widetilde{\mathcal{K}}} \in \mathcal{G}(\boldsymbol{y})\right\}.$$

Then, from Lemma 9, we have the fact that $\boldsymbol{x}_{\mathcal{J}^c} \in \mathcal{G}_{\mathcal{X}_{\mathcal{J}^c}}(\boldsymbol{y})$ implies $(\boldsymbol{x}_{\mathcal{J}^c}, \boldsymbol{y}) \in \mathcal{T}_{X_{\mathcal{J}^c}Y,\gamma'}$ and

$$\mathcal{G}_{\mathcal{X}_{\mathcal{J}}|\mathcal{X}_{\mathcal{J}^c}}(\boldsymbol{x}_{\mathcal{J}^c}, \boldsymbol{y}) \subset \mathcal{T}_{X_{\mathcal{J}}|X_{\mathcal{J}^c}Y,\gamma'}(\boldsymbol{x}_{\mathcal{J}^c}, \boldsymbol{y})$$

for every non-empty set $\mathcal{J} \subsetneq \widetilde{\mathcal{K}}$. We have

$$\begin{aligned}
|\mathcal{G}_{\mathcal{J}|\mathcal{J}^c}(\boldsymbol{y})| &\equiv \max_{\boldsymbol{x}_{\mathcal{J}^c} \in \mathcal{G}_{\mathcal{X}_{\mathcal{J}^c}}(\boldsymbol{y})} \left|\mathcal{G}_{\mathcal{X}_{\mathcal{J}}|\mathcal{X}_{\mathcal{J}^c}}(\boldsymbol{x}_{\mathcal{J}^c}, \boldsymbol{y})\right| \\
&\leq \max_{(\boldsymbol{x}_{\mathcal{J}^c}, \boldsymbol{y}) \in \mathcal{T}_{X_{\mathcal{J}^c} Y, \gamma'}} \left|\mathcal{T}_{X_{\mathcal{J}}|X_{\mathcal{J}^c} Y, \gamma'}(\boldsymbol{x}_{\mathcal{J}^c}, \boldsymbol{y})\right| \\
&\leq 2^{n[H(X_{\mathcal{J}}|X_{\mathcal{J}^c}, Y) + \eta_{\mathcal{X}_{\mathcal{J}}|\mathcal{X}_{\mathcal{J}^c}\mathcal{Y}}(\gamma'|\gamma')]} \\
&\leq 2^{n[H(X_{\mathcal{J}}|X_{\mathcal{J}^c}, Y) + \eta_{\mathcal{X}_{\widetilde{\mathcal{K}}}|\mathcal{Y}}(\gamma'|\gamma')]}
\end{aligned} \tag{126}$$

for every non-empty set $\mathcal{J} \subsetneq \widetilde{\mathcal{K}}$, where the second inequality comes from Lemma 13. Then, from (125), (126), and Lemma 4, we have

$$\begin{aligned}
E_{\mathsf{A}_{\widetilde{\mathcal{K}}}}\left[\chi(\widehat{g}_{\mathsf{A}_{\widetilde{\mathcal{K}}}}(\mathsf{A}_{\widetilde{\mathcal{K}}}\boldsymbol{x}_{\widetilde{\mathcal{K}}}|\boldsymbol{y}) \neq \boldsymbol{x}_{\widetilde{\mathcal{K}}})\right] &\leq p_{\mathsf{A}_{\widetilde{\mathcal{K}}}}\left(\left\{A_{\widetilde{\mathcal{K}}} : \left[\mathcal{G}(\boldsymbol{y}) \setminus \{\boldsymbol{x}_{\widetilde{\mathcal{K}}}\}\right] \cap \mathcal{C}_{A_{\widetilde{\mathcal{K}}}}(A_{\widetilde{\mathcal{K}}}\boldsymbol{x}_{\widetilde{\mathcal{K}}}) \neq \emptyset\right\}\right) \\
&\leq \sum_{\substack{\mathcal{J} \subset \widetilde{\mathcal{K}} \\ \mathcal{J} \neq \emptyset}} \frac{2^{n\left[H(X_{\mathcal{J}}|X_{\mathcal{J}^c}, Y) + \eta_{\mathcal{X}_{\widetilde{\mathcal{K}}}|\mathcal{Y}}(\gamma'|\gamma')\right]} \alpha_{\mathsf{A}_{\mathcal{J}}}\left[\beta_{\mathsf{A}_{\mathcal{J}^c}} + 1\right]}{\prod_{j \in \mathcal{J}} |\mathrm{Im}\mathcal{A}_j|} + \beta_{\mathsf{A}_{\widetilde{\mathcal{K}}}}
\end{aligned} \tag{127}$$

for all $(\boldsymbol{x}_{\widetilde{\mathcal{K}}}, \boldsymbol{y}) \in \mathcal{T}_{X_{\widetilde{\mathcal{K}}} Y, \gamma'}$. Then we have

$$\begin{aligned}
&E_{\widehat{\mathsf{A}}_{\widetilde{\mathcal{K}}}\mathbf{a}_{\widetilde{\mathcal{K}}}}\left[p_{M_{\widetilde{\mathcal{K}}}Y}\left(\left[\cap_{j=0}^{4}\mathcal{S}_j\right] \cap \mathcal{S}_5^c\right)\right] \\
&\leq E_{\widehat{\mathsf{A}}_{\widetilde{\mathcal{K}}}\widehat{\mathbf{a}}_{\widetilde{\mathcal{K}}}}\left[\sum_{\boldsymbol{x}_{\widetilde{\mathcal{K}}} \in \mathcal{T}} \chi(\widehat{g}_{\widehat{\mathsf{A}}_0}(\widehat{\mathbf{a}}_0) = \boldsymbol{x}_0)\left[\prod_{j \in \mathcal{K}} \chi(\widehat{g}_{\widehat{\mathsf{A}}_j}(\widehat{\mathbf{a}}_j|\boldsymbol{x}_0) = \boldsymbol{x}_j)\right] \sum_{\boldsymbol{y} \in \mathcal{T}_{Y|X_{\widetilde{\mathcal{K}}},\gamma}(\boldsymbol{x}_{\widetilde{\mathcal{K}}})} \mu_{Y|X_{\widetilde{\mathcal{K}}}}(\boldsymbol{y}|\boldsymbol{x}_{\widetilde{\mathcal{K}}})\chi(\widehat{g}_{\mathsf{A}_{\widetilde{\mathcal{K}}}}(\mathbf{a}_{\widetilde{\mathcal{K}}}|\boldsymbol{y}) \neq \boldsymbol{x}_{\widetilde{\mathcal{K}}})\right] \\
&\leq \sum_{\substack{\boldsymbol{x}_{\widetilde{\mathcal{K}}} \in \mathcal{T} \\ \boldsymbol{y} \in \mathcal{T}_{Y|X_{\widetilde{\mathcal{K}}},\gamma}(\boldsymbol{x}_{\widetilde{\mathcal{K}}})}} \mu_{Y|X_{\widetilde{\mathcal{K}}}}(\boldsymbol{y}|\boldsymbol{x}_{\widetilde{\mathcal{K}}}) E_{\mathsf{A}_{\widetilde{\mathcal{K}}}}\left[\chi(\widehat{g}_{\mathsf{A}_{\widetilde{\mathcal{K}}}}(\mathsf{A}_{\widetilde{\mathcal{K}}}\boldsymbol{x}_{\widetilde{\mathcal{K}}}|\boldsymbol{y}) \neq \boldsymbol{x}_{\widetilde{\mathcal{K}}}) \prod_{j \in \widetilde{\mathcal{K}}} E_{\mathbf{a}_j}\left[\chi(\mathsf{A}_j\boldsymbol{x}_j = \mathbf{a}_j)\right] E_{\mathsf{A}'_j M_j}\left[\chi(\mathsf{A}'_j\boldsymbol{x}_j = M_j)\right]\right] \\
&= \frac{1}{\prod_{j \in \widetilde{\mathcal{K}}} |\mathrm{Im}\widehat{A}_j|} \sum_{\substack{\boldsymbol{x}_{\widetilde{\mathcal{K}}} \in \mathcal{T} \\ \boldsymbol{y} \in \mathcal{T}_{Y|X_{\widetilde{\mathcal{K}}},\gamma}(\boldsymbol{x}_{\widetilde{\mathcal{K}}})}} \mu_{Y|X_{\widetilde{\mathcal{K}}}}(\boldsymbol{y}|\boldsymbol{x}_{\widetilde{\mathcal{K}}}) E_{\mathsf{A}_{\widetilde{\mathcal{K}}}}\left[\chi(\widehat{g}_{\mathsf{A}_{\widetilde{\mathcal{K}}}}(\mathsf{A}_{\widetilde{\mathcal{K}}}\boldsymbol{x}_{\widetilde{\mathcal{K}}}|\boldsymbol{y}) \neq \boldsymbol{x}_{\widetilde{\mathcal{K}}})\right] \\
&\leq \frac{1}{\prod_{j \in \widetilde{\mathcal{K}}} |\mathrm{Im}\widehat{A}_j|} \sum_{\substack{\boldsymbol{x}_{\widetilde{\mathcal{K}}} \in \mathcal{T} \\ \boldsymbol{y} \in \mathcal{T}_{Y|X_{\widetilde{\mathcal{K}}},\gamma}(\boldsymbol{x}_{\widetilde{\mathcal{K}}})}} \mu_{Y|X_{\widetilde{\mathcal{K}}}}(\boldsymbol{y}|\boldsymbol{x}_{\widetilde{\mathcal{K}}}) \left[\sum_{\substack{\mathcal{J} \subset \widetilde{\mathcal{K}} \\ \mathcal{J} \neq \emptyset}} \frac{2^{n\left[H(X_{\mathcal{J}}|X_{\mathcal{J}^c}, Y) + \eta_{\mathcal{X}_{\widetilde{\mathcal{K}}}|\mathcal{Y}}(\gamma'|\gamma')\right]} \alpha_{\mathsf{A}_{\mathcal{J}}}\left[\beta_{\mathsf{A}_{\mathcal{J}^c}} + 1\right]}{\prod_{j \in \mathcal{J}} |\mathrm{Im}\mathcal{A}_j|} + \beta_{\mathsf{A}_{\widetilde{\mathcal{K}}}}\right] \\
&\leq 8\kappa^3 \left[\sum_{\substack{\mathcal{J} \subset \widetilde{\mathcal{K}} \\ \mathcal{J} \neq \emptyset}} 2^{-n\left[\sum_{j \in \mathcal{J}} r_j - H(X_{\mathcal{J}}|X_{\mathcal{J}^c}, Y) - \eta_{\mathcal{X}_{\widetilde{\mathcal{K}}}|\mathcal{Y}}(\gamma'|\gamma')\right]} \alpha_{\mathsf{A}_{\mathcal{J}}}\left[\beta_{\mathsf{A}_{\mathcal{J}^c}} + 1\right] + \beta_{\mathsf{A}_{\widetilde{\mathcal{K}}}}\right] \\
&\leq \frac{\delta}{6}
\end{aligned} \tag{128}$$

for all $\delta > 0$ and all sufficiently large n, where $\mathcal{T}$ is defined as

$$\mathcal{T} \equiv \{(\boldsymbol{x}_0, \boldsymbol{x}_1, \boldsymbol{x}_2) : \boldsymbol{x}_0 \in \mathcal{T}_0, \boldsymbol{x}_1 \in \mathcal{T}_1(\boldsymbol{x}_0), \boldsymbol{x}_2 \in \mathcal{T}_2(\boldsymbol{x}_0)\},$$

the equality comes from Lemma 5, which appears in Appendix A, the third inequality comes from (127), the fourth inequality comes from (54), (110), and (111), and the last inequality comes from (26), (27), (59)–(65), and (105).

Finally, from (113)–(120), (122), (123), and (128), we have the fact that for all $\delta > 0$ and sufficiently large n there are $\{A_j, A'_j, \boldsymbol{a}_j\}_{j \in \widetilde{\mathcal{K}}}$ satisfying $A_j \in \mathcal{A}$, $A'_j \in \mathcal{A}'_j$, $\boldsymbol{a}_j \in \mathrm{Im}\mathcal{A}_j$, and (66). ■

APPENDIX

A. Basic Property of Ensemble

Lemma 5 ([19, Lemma 9]): Assume that random variables A and $\mathbf{a}$ are independent and the distribution of $\mathbf{a}$ is uniform on $\mathrm{Im}\mathcal{A}$. Then,

$$E_{\mathbf{a}}\left[\chi(A\boldsymbol{u}=\mathbf{a})\right]=\frac{1}{|\mathrm{Im}\mathcal{A}|}$$

for any $A\in\mathcal{A}$ and $\boldsymbol{u}\in\mathcal{U}^n$, and

$$E_{\mathsf{A}\mathbf{a}}\left[\chi(\mathsf{A}\boldsymbol{u}=\mathbf{a})\right]=\frac{1}{|\mathrm{Im}\mathcal{A}|}$$

for any $\boldsymbol{u}\in\mathcal{U}^n$.

B. Method of Types

Let $\mathcal{T}_U\subset\mathcal{U}^n$ be a set of all sequences that has the same type ν_U, where type of $\boldsymbol{u}\in\mathcal{U}^n$ is defined by the empirical distribution $\nu_{\boldsymbol{u}}$. Let $\mathcal{T}_{U,\gamma}$ be a set of typical sequences and $\mathcal{T}_{U|V,\gamma}(\boldsymbol{v})$ be a set of conditionally typical sequences defined in the beginning of Section II.

Lemma 6 ([8, Lemma 2.2]): The number of different types of sequences in $\mathcal{U}^n$ is fewer than $[n+1]^{|\mathcal{U}|}$. The number of conditional types of sequences in $\mathcal{U}^n\times\mathcal{V}^n$ is fewer than $[n+1]^{|\mathcal{U}||\mathcal{V}|}$.

Lemma 7 ([8, Lemma 2.3 and 2.5]): Let $\lambda_{\mathcal{U}}$ be defined in (3). Then

$$2^{n[H(U)-\lambda_{\mathcal{U}}]}\le|\mathcal{T}_U|\le 2^{nH(U)}.$$

Lemma 8: For $H\ge 0$,

$$|\,\{\boldsymbol{u}:H(\boldsymbol{u})\le H\}\,|\le 2^{n[H+\lambda_{\mathcal{U}}]}$$

where $\lambda_{\mathcal{U}}$ is defined by (3).

Proof: The proof is similar to that of [20, Lemma 6]. We have

$$\begin{aligned}|\,\{\boldsymbol{u}:H(\boldsymbol{u})\le H\}\,| &= \sum_{U:H(U)\le H}|\mathcal{T}_U|\\ &\le \sum_{U:H(U)\le H}2^{nH(U)}\\ &\le \sum_{U:H(U)\le H}2^{nH}\\ &\le [n+1]^{|\mathcal{U}|}2^{nH}\\ &= 2^{n[H+\lambda_{\mathcal{U}}]},\end{aligned}\tag{129}$$

where the sum is taken over all random variables U corresponding the type of a sequence in $\mathcal{U}^n$, the first inequality comes from Lemma 7, and the last inequality comes from Lemma 6. ■

Lemma 9 ([19, Lemma 22][27, Theorem 2.5]): If $\boldsymbol{v}\in\mathcal{T}_{V,\gamma}$ and $\boldsymbol{u}\in\mathcal{T}_{U|V,\gamma'}(\boldsymbol{v})$, then $(\boldsymbol{u},\boldsymbol{v})\in\mathcal{T}_{UV,\gamma+\gamma'}$. If $(\boldsymbol{u},\boldsymbol{v})\in\mathcal{T}_{UV,\gamma}$, then $\boldsymbol{u}\in\mathcal{T}_{U,\gamma}$ and $\boldsymbol{u}\in\mathcal{T}_{U|V,\gamma}(\boldsymbol{v})$.

Lemma 10 ([27, Theorem 2.6]): If $\boldsymbol{u}\in\mathcal{T}_{U,\gamma}$, then

$$\sum_{u\in\mathcal{U}}|\nu_{\boldsymbol{u}}(u)-\mu_U(u)|\le\sqrt{2\gamma}$$

Proof: The statement is shown by

$$\begin{aligned}\sum_{u\in\mathcal{U}}|\nu_{\boldsymbol{u}}(u)-\mu_U(u)| &\le \sqrt{\frac{2D(\nu_{\boldsymbol{u}}\parallel\mu_U)}{\log_2 e}}\\ &\le \sqrt{\frac{2\gamma}{\log_2 e}}\\ &\le \sqrt{2\gamma}, \end{aligned} \tag{130}$$

where e is the base of the natural logarithm and the first inequality comes from [6, Lemma 12.6.1]. ■

Lemma 11: Let $0<\gamma\le 1/8$. If $\boldsymbol{v}\in\mathcal{T}_{V,\gamma}$, and $\boldsymbol{u}\in\mathcal{T}_{U|V,\gamma'}(\boldsymbol{v})$, then

$$|H(\boldsymbol{v})-H(V)|\le \iota_{\mathcal{V}}(\gamma)$$
$$|H(\boldsymbol{u}|\boldsymbol{v})-H(U|V)|\le \iota_{\mathcal{U}|\mathcal{V}}(\gamma'|\gamma),$$

where $\iota_{\mathcal{U}}$ and $\iota_{\mathcal{U}|\mathcal{V}}$ are defined by (4) and (5), respectively.

Proof: From [8, Lemma 2.7], we have

$$|H(p)-H(q)|\le -\theta\log\frac{\theta}{|\mathcal{U}|} \tag{131}$$

for any θ and probability distributions p and q on $\mathcal{V}$ satisfying

$$\sum_{v\in\mathcal{V}}|p(v)-q(v)|\le\theta\le\frac{1}{2}.$$

Then the first inequality is shown by this fact and Lemma 10.

Next we prove the second inequality. Let $\nu_{\boldsymbol{u}|\boldsymbol{v}}\nu_{\boldsymbol{v}}$ and $\mu_{U|V}\nu_{\boldsymbol{v}}$ be defined as

$$\nu_{\boldsymbol{u}|\boldsymbol{v}}\nu_{\boldsymbol{v}}(u,v)\equiv\nu_{\boldsymbol{u}|\boldsymbol{v}}(u|v)\nu_{\boldsymbol{v}}(v)$$
$$\mu_{U|V}\nu_{\boldsymbol{v}}(u,v)\equiv\mu_{U|V}(u|v)\nu_{\boldsymbol{v}}(v),$$

respectively. Since

$$\begin{aligned}D(\nu_{\boldsymbol{u}|\boldsymbol{v}}\nu_{\boldsymbol{v}}\parallel\mu_{U|V}\nu_{\boldsymbol{v}}) &= \sum_{u,v}\nu_{\boldsymbol{u}|\boldsymbol{v}}(u|v)\nu_{\boldsymbol{v}}(v)\log\frac{\nu_{\boldsymbol{u}|\boldsymbol{v}}(u|v)}{\mu_{U|V}(u|v)}\\ &= D(\nu_{\boldsymbol{u}|\boldsymbol{v}}\parallel\mu_{U|V}|\nu_{\boldsymbol{v}})\\ &< \gamma'\end{aligned} \tag{132}$$

we have

$$\begin{aligned}&|H(\nu_{\boldsymbol{u}|\boldsymbol{v}}|\nu_{\boldsymbol{v}})-H(\mu_{U|V}|\nu_{\boldsymbol{v}})|\\ &=\left|\sum_{u,v}\nu_{\boldsymbol{u}|\boldsymbol{v}}(u|v)\nu_{\boldsymbol{v}}(v)\log\frac{1}{\nu_{\boldsymbol{u}|\boldsymbol{v}}(u|v)}-\sum_{u,v}\mu_{U|V}(u|v)\nu_{\boldsymbol{v}}(v)\log\frac{1}{\mu_{U|V}(u|v)}\right|\\ &=\left|\sum_{u,v}\nu_{\boldsymbol{u}|\boldsymbol{v}}(u|v)\nu_{\boldsymbol{v}}(v)\log\frac{\nu_{\boldsymbol{v}}(v)}{\nu_{\boldsymbol{u}|\boldsymbol{v}}(u|v)\nu_{\boldsymbol{v}}(v)}-\sum_{u,v}\mu_{U|V}(u|v)\nu_{\boldsymbol{v}}(v)\log\frac{\nu_{\boldsymbol{v}}(v)}{\mu_{U|V}(u|v)\nu_{\boldsymbol{v}}(v)}\right|\\ &=|H(\nu_{\boldsymbol{u}|\boldsymbol{v}}\nu_{\boldsymbol{v}})-H(\mu_{U|V}\nu_{\boldsymbol{v}})|\\ &\le\iota_{\mathcal{U}\mathcal{V}}(\gamma'),\end{aligned} \tag{133}$$

where the last inequality comes from (131). We have

$$\begin{aligned}|H(\nu_{\boldsymbol{u}|\boldsymbol{v}}|\nu_{\boldsymbol{v}}) - H(\nu_{\boldsymbol{u}|\boldsymbol{v}}|\mu_V)| &= \left|\sum_{u,v}\nu_{\boldsymbol{u}|\boldsymbol{v}}(u|v)\nu_{\boldsymbol{v}}(v)\log\frac{1}{\nu_{\boldsymbol{u}|\boldsymbol{v}}(u|v)} - \sum_{u,v}\nu_{\boldsymbol{u}|\boldsymbol{v}}(u|v)\mu_V(v)\frac{1}{\nu_{\boldsymbol{u}|\boldsymbol{v}}(u|v)}\right| \\ &\le \sum_v |\nu_{\boldsymbol{v}}(v) - \mu_V(v)| \sum_u \nu_{\boldsymbol{u}|\boldsymbol{v}}(u|v)\log\frac{1}{\nu_{\boldsymbol{u}|\boldsymbol{v}}(u|v)} \\ &= \sum_v |\nu_{\boldsymbol{v}}(v) - \mu_V(v)| H(\nu_{\boldsymbol{u}|\boldsymbol{v}}(\cdot|v)) \\ &\le \sqrt{2\gamma}\log|\mathcal{U}|, \end{aligned} \tag{134}$$

where the last inequality comes from Lemma 10 and the fact that $H(\nu_{\boldsymbol{u}|\boldsymbol{v}}(\cdot|v)) \le \log|\mathcal{U}|$. From (133) and (134), we have

$$\begin{aligned}|H(\boldsymbol{u}|\boldsymbol{v}) - H(U|V)| &\le |H(\nu_{\boldsymbol{u}|\boldsymbol{v}}|\nu_{\boldsymbol{v}}) - H(\mu_{U|V}|\nu_{\boldsymbol{v}})| + |H(\mu_{U|V}|\nu_{\boldsymbol{v}}) - H(U|V)| \\ &\le \iota_{\mathcal{U}|\mathcal{V}}(\gamma'|\gamma). \end{aligned} \tag{135}$$

■

Lemma 12 ([19, Lemma 26][27, Theorem 2.8]): For any $\gamma > 0$, and $\boldsymbol{v} \in \mathcal{V}^n$,

$$\mu_U([\mathcal{T}_{U,\gamma}]^c) \le 2^{-n[\gamma-\lambda_{\mathcal{U}}]}$$

$$\mu_{U|V}([\mathcal{T}_{U|V,\gamma}(\boldsymbol{v})]^c|\boldsymbol{v}) \le 2^{-n[\gamma-\lambda_{\mathcal{U}\mathcal{V}}]},$$

where $\lambda_{\mathcal{U}}$ and $\lambda_{\mathcal{U}\mathcal{V}}$ are defined in (3).

Lemma 13 ([19, Lemma 27][27, Theorem 2.9]): For any $\gamma > 0$, $\gamma' > 0$, and $\boldsymbol{v} \in \mathcal{T}_{V,\gamma}$,

$$\left|\frac{1}{n}\log|\mathcal{T}_{U,\gamma}| - H(U)\right| \le \eta_{\mathcal{U}}(\gamma)$$

$$\left|\frac{1}{n}\log|\mathcal{T}_{U|V,\gamma'}(\boldsymbol{v})| - H(U|V)\right| \le \eta_{\mathcal{U}|\mathcal{V}}(\gamma'|\gamma),$$

where $\eta_{\mathcal{U}}(\gamma)$ and $\eta_{\mathcal{U}|\mathcal{V}}(\gamma'|\gamma)$ are defined in (6) and (7), respectively.

ACKNOWLEDGEMENTS

We thank Prof. T.S. Han for helpful discussions. Constructive comments, suggestions, and references by anonymous reviewers have significantly improved the presentation of our results.

REFERENCES

[1] R. Ahlswede, "Multi-way communication channels," *Proc. 2nd International Symposium on Information Theory,* Tsahkadsor Armenian SSR, pp. 23–52, 1971.

[2] A. Amaraoui, S. Dusad, R. Urbanke, "Achieving general points in the 2-user Gaussian MAC without time-sharing or rate-splitting by means of iterative coding," *Proc. 2002 International Symposium on Information Theory,* Lausanne, Switzerland, Jun. 30–Jul. 5, 2002, p. 334.

[3] A. de Baynast and D. Declercq, "Gallager codes for multiple user applications," *Proc. 2002 International Symposium on Information Theory,* Lausanne, Switzerland, Jun. 30–Jul. 5, 2002, p. 335.

[4] T. M. Cover, "Broadcast Channels," *IEEE Trans. Inform Theory*, vol. IT-18, no. 1, pp. 2–14, Jan. 1972.

[5] T. M. Cover, "A proof of the data compression theorem of Slepian and Wolf for ergodic sources," *IEEE Trans. Inform Theory*, vol. IT-21, no. 2, pp. 226–228, Mar. 1975.

[6] T. M. Cover and J. A. Thomas, *Elements of Information Theory 2nd. Ed.*, John Wiley & Sons, Inc., 2006.

[7] I. Csiszár, “Linear codes for sources and source networks: Error exponents, universal coding,” *IEEE Trans. Inform. Theory*, vol. IT-28, no. 4, pp. 585–592, Jul. 1982.

[8] I. Csiszár and J. Körner, *Information Theory: Coding Theorems for Discrete Memoryless Systems*, Academic Press, 1981.

[9] J. L. Carter and M. N. Wegman, “Universal classes of hash functions,” *J. Comput. Syst. Sci.*, vol. 18, pp. 143–154, 1979.

[10] A. El Gamal and Y.H. Kim, *Network information theory,* Cambridge University Press, 2011.

[11] J. Feldman, M.J. Wainwright, and D.R. Karger, “Using linear programming to decode binary linear codes,” *IEEE Trans. Inform. Theory*, vol. IT-51, no. 3, pp. 954–972, Mar. 2005.

[12] T.S. Han, “The capacity region of general multiple-access channel with certain correlated sources,” *Inform. Contr.*, vol.40, pp.37–60, 1979.

[13] T.S. Han and K. Kobayashi, “A new achievable rate region for the interference channel,” *IEEE Trans. Info. Theory*, vol. IT-27, no. 1, pp. 49–60, Jan. 1981.

[14] F. R. Kschischang, B. J. Frey, and H. A. Loeliger, “Factor graphs and the sum-product algorithm,” *IEEE Transactions on Information Theory*, vol. 47, no. 2, pp. 498–519, Feb. 2001.

[15] H. Liao, *Multiple Access Channels*, Ph.D. thesis, Department of Electrical Engineering, University of Hawaii, Honolulu, 1972.

[16] R. J. McEliece, “Are Turbo-like codes effective on nonstandard channels?” *IEEE Information Theory Society Newsletter,* vol. 51, no. 4, p. 1 and pp.3–8, 2001.

[17] J. Muramatsu, T. Uyematsu, and T. Wadayama, “Low density parity check matrices for coding of multiple access networks,” *Proc. IEEE Information Theory Workshop*, Paris, France, Mar. 31–Apr. 4, 2003, pp. 304–307.

[18] J. Muramatsu, T. Uyematsu, and T. Wadayama, “Low density parity check matrices for coding of correlated sources,” *IEEE Trans. Inform. Theory*, vol. IT-51, no. 10, pp. 3645–3653, Oct. 2005.

[19] J. Muramatsu and S. Miyake, “Hash property and coding theorems for sparse matrices and maximal-likelihood coding,” *IEEE Trans. Inform. Theory*, vol. IT-56, no. 5, pp. 2143–2167, May 2010. Corrections: vol. IT-56, no. 9, p. 4762, Sept. 2010.

[20] J. Muramatsu and S. Miyake, “Hash property and fixed-rate universal coding theorems,” *IEEE Trans. Inform. Theory*, vol. IT-56, no. 6, pp. 2688–2698, Jun. 2010. Corrections: vol. IT-58, no. 5, pp. 3305–3307, May 2012.

[21] J. Muramatsu and S. Miyake “Construction of Slepian-Wolf source code and broadcast channel code based on hash property,” available at `arXiv:1006.5271[cs.IT]`, 2010.

[22] J. Muramatsu and S. Miyake, “Construction of broadcast channel code based on hash property,” *Proc. 2010 IEEE Int. Symp. Inform. Theory*, Austin, U.S.A., June 13–18, 2010, pp. 575–579.

[23] J. Muramatsu and S. Miyake, “Construction of strongly secure wiretap channel code based on hash property,” *Proc. of 2011 IEEE Int. Symp. Inform. Theory*, St. Petersburg, Russia, Jul. 31–Aug. 5, 2011, pp.612–616.

[24] J. Muramatsu and S. Miyake, “Construction of multiple-access channel codes based on hash property,” *Proc. of 2011 IEEE Int. Symp. Inform. Theory*, St. Petersburg, Russia, Jul. 31–Aug. 5, 2011, pp.2274–2278.

[25] A. Sanderovich, M. Peleg, and S. Shamai, “LDPC coded MIMO multiple access with iterative joint decoding,” *IEEE Trans. Inform. Theory*, vol.IT-51, no. 4, pp.1437–1450, Apr. 2005.

[26] D. Slepian and J. K. Wolf, “A coding theorem for multiple access channels with correlated sources,” *Bell System Technical Journal*, vol. 52, no. 7, pp. 1037–1076, Sep. 1973.

[27] T. Uyematsu, *Gendai Shannon Riron*, Baifukan, 1998 (in Japanese).

[28] G. M. Ziegler, *Lectures on Polytopes*, Springer Science+Business Media, LLC, 2006.